
Autocorrelation Studies of the Arrival Directions of UHECRs measured by the Surface Detector of the Pierre Auger Observatory

Von der Fakultät für Mathematik, Informatik und Naturwissenschaften der RWTH
Aachen University zur Erlangung des akademischen Grades eines Doktors der
Naturwissenschaften genehmigte Dissertation

vorgelegt von

Diplom-Physiker

Stephan Schulte

aus Arnsberg (Westfalen)

Berichter: Universitätsprofessor Thomas Hebbeker

Universitätsprofessor Martin Erdmann

Tag der mündlichen Prüfung: 11.07.2011

Diese Dissertation ist auf den Internetseiten der Hochschulbibliothek online verfü-
bar.

Erstgutachter und Betreuer

Prof. Dr. Thomas Hebbeker
III. Physikalisches Institut A
RWTH Aachen

Zweitgutachter

Prof. Dr. Martin Erdmann
III. Physikalisches Institut A
RWTH Aachen

Contents

Outline	1
1 Introduction	3
2 Coordinate Systems	5
2.1 Local Coordinates (LCS)	5
2.2 Equatorial Coordinates (ECS)	5
2.3 Galactic Coordinates (GCS)	6
2.4 Super Galactic Coordinates (SGCS)	6
2.5 Transformation	6
3 Cosmic Rays	9
3.1 Fermi Acceleration	11
3.1.1 Second Order Fermi Acceleration	12
3.1.2 First Order Fermi Acceleration	13
3.2 The Origin of Cosmic Rays up to the knee	14
3.3 The Origin of UHECRs	14
3.3.1 Bottom-Up Models	16
3.3.1.1 Active Galactic Nuclei	16
3.3.1.2 Gamma Ray Bursts	21
3.3.2 Top-Down Models	21
3.3.3 Propagation	21
3.3.3.1 GZK-Effect	23
3.3.3.2 Magnetic Fields	23
3.3.4 Low Energy Anisotropy Studies	25
3.3.5 Autocorrelation Studies	26
3.4 Extensive Air Showers	27

3.4.1	Shower Development	28
3.4.1.1	Electromagnetic Shower	28
3.4.1.2	Hadronic Showers	29
3.4.2	Detection Techniques of EAS	30
3.4.2.1	Measuring the Longitudinal Shower Profile	31
3.4.2.2	Measuring the Lateral Shower Profile	33
4	The Pierre Auger Observatory	37
4.1	Fluorescence Detector	38
4.1.1	Reconstruction Principle and Resolution	39
4.2	Surface Detector	41
4.2.1	Trigger Structure	41
4.2.2	Reconstruction and Resolution	44
4.3	Enhancements	46
4.3.1	High Elevation Auger Telescopes (HEAT)	49
4.3.1.1	TILT-Monitoring	50
4.3.1.2	TILT-Setup	55
4.3.2	Auger Muons and Infill for the Ground Array (AMIGA)	57
4.3.3	Auger Engineering Radio Array (AERA)	58
5	Autocorrelation Methods	59
5.1	2pt-Correlation-Function	60
5.1.1	Using one angle	60
5.1.2	Using two angles	64
5.2	Minimum Spanning Tree	65
5.3	Cluster Algorithm	68
5.3.1	Unweighted Realisation (CA I)	70
5.3.2	Weighted Realisation (CA II)	73
5.4	Summary	73

6	Monte Carlo Studies	77
6.1	Generating Mock Maps	77
6.1.1	HEALPix	78
6.1.2	Smearing	78
6.1.3	Effect of Coverage	80
6.1.4	Catalogues	80
6.1.4.1	VCV	80
6.1.4.2	IRAS	82
6.1.5	Single Sources	83
6.1.5.1	Radio Loud Galaxies	83
6.1.5.2	Random Sources	87
6.1.6	Multipoles	87
6.2	Analysing Mock Maps	88
6.2.1	Analysis of catalogue-based Mock Maps	90
6.2.2	Analysis of single-source Mock Maps	92
6.2.3	Analysis of Multipoles	96
6.3	Summary and Conclusions	97
7	Application to Data measured with the PAO	101
7.1	Cut Parameters	101
7.2	High Energy Data Sets	102
7.2.1	Comparison with the result of the AGN correlation study . . .	103
7.2.2	Current Data Set	103
7.2.2.1	Comparison between Herald and Observer	104
7.3	Low Energy Data Set	105
7.4	Point Source Search	107
7.5	Summary	109
8	Conclusion & Outlook	113
A	Acronyms	117
A.1	List of abbreviations	117

B Tilt Monitoring System	119
B.1 Calibration Constants (Tilt-Monitoring Setup)	119
B.2 Software Screenshots	120
C Significance Calculation	123
C.1 Significance Calculation in the case of combined tests	123
D Monte Carlo Maps	125
D.1 Produced Monte Carlo Maps	125
E Optimal Radius of the Cluster Algorithm	129
E.1 Dependence of the Cluster Algorithm on the Radius	129
E.1.1 Summary	133
F Results of the Monte Carlo Study	135
F.1 Maps based on catalogues	135
F.2 Maps based on single sources	143
F.3 Maps based on multipoles	143
F.4 Summary	150
References	156
Acknowledgements	157

List of Figures

2.1	Super Galactic Coordinates	7
3.1	Cosmic Ray Flux	10
3.2	X_{max} -Distribution fmeasure by the PAO	11
3.3	Flux of Cosmic Rays at the highest energies	15
3.4	Hillas plot and Loss Length for some nuclei	16
3.5	AGN-Scheme	17
3.6	AGN-Correlation result of the PAO	18
3.7	Cumulative Distance Distribution around CenA	20
3.8	Photon and τ -Neutrino Limits of PAO	22
3.9	Simple Propagation Simulation including Magnetic Fields	24
3.10	Right Ascension Modulation Analysis by PAO	27
3.11	Scheme of the Heitler model	31
3.12	Scheme of Detection principles of Extensive Air Showers	32
3.13	Example PMT track measured by a fluorescence telescope of the PAO	33
3.14	Gaisser Hillas Fit to an event measured by PAO	34
3.15	NKG Fit to an event measured by the PAO	35
4.1	Map of the Pierre Auger Observatory	38
4.2	The Fluorescence Detector of PAO	39
4.3	Allowed shapes of the second level Trigger of FD	40
4.4	Geometrical shower description	41
4.5	An example hybrid FD event	42
4.6	Surface Detector Station	42
4.7	SD T3 trigger configuration	44
4.8	SD Trigger efficiency	45
4.9	Example Footprint of an SD Event	47

4.10	Example NKG function of an SD Event	47
4.11	Correlation between FD energy and SD energy estimator	48
4.12	Example Timing fit of an SD Event	48
4.13	Site of the Auger Enhancements	49
4.14	Photo of all three HEAT telescopes inclined	50
4.15	HEAT and Coihueco Stereo Event	51
4.16	Tilt-Monitoring Sensor Positions	53
4.17	Sensors of the Tilt-Setup	53
4.18	Calibration Setup of the inclination sensors of the Tilt-Setup	54
4.19	Difference between measurements and true value for the calibrations sensors of the Tilt-Setup	55
4.20	Residuum of calibration of the inclination sensors for the Tilt-Setup .	56
4.21	Scheme of the current TILT-Setup	57
4.22	A typical AMIGA station	58
4.23	The layout of AERA	58
5.1	Example MC Data Set based on the VCV	60
5.2	Solid Angle Distribution for Test Data Set	61
5.3	Example Test Set:TwoPt Cumulative Distance Distribution	63
5.4	Example Test Set: TwoPt Estimators	63
5.5	Example Test Set:TwoPt Combination of Estimators	64
5.6	Example Test Set: TwoPtPlus	65
5.7	Example Test Set: TwoPtPlus Estimator Distributions	66
5.8	Minimum Spanning Tree Flow Chart	67
5.9	Minimum Spanning Tree Example	68
5.10	Example Set analysed according to the Minimum Spanning Tree . . .	69
5.11	Example Test Set: Minimum Spanning Tree	69
5.12	Example Data Set: Cluster Algorithm	71
5.13	Simple Cluster Algorithm Examples	72
5.14	Results of the Cluster Algorithm applied to a test data set	72
5.15	Distribution of Cluster Size for 10,000 isotropic sets of 100 events. Doublets with small weights dominate.	74

5.16	Distribution of the Cluster Sum and the test data set (CA II)	75
6.1	HEALPix scheme for large pixel sizes	79
6.2	The Coverage of the Pierre Auger Observatory	81
6.3	The Veron-Cetty and Veron catalogue	82
6.4	The Infrared Astronomical Satellite catalogue	84
6.5	Probability Density Maps based on three radio loud galaxies	86
6.6	Distribution around Source corresponding to the Inverse model	87
6.7	Example Monte Carlo Set based on the ellipse model	88
6.8	Simulated Multipoles according to the coverage of the PAO	89
6.9	Sensitivity of Autocorrelation methods applied to MC maps based on catalogues	90
6.10	Sensitivity of Autocorrelation methods applied to MC maps based on radio loud galaxies	94
6.11	Sensitivity of Autocorrelation methods applied to MC maps based on the ellipse model	96
6.12	Sensitivity of Autocorrelation methods applied to MC maps based on multipoles	97
7.1	CA II analysis of published data	104
7.2	Comparison of the two reconstruction algorithms	105
7.3	Cluster Sum distribution for the high energy sets	106
7.4	Probability of the current high energy data set for Herald and Observer	107
7.5	Cluster Size distribution for the low energy data set	108
7.6	Cluster Sum distribution for the low energy data set	108
7.7	Point Source analysis with CA II regarding Cen A	110
B.1	Screenshots of the TILT-Monitoring Software	121
C.1	Probability Correction for combined tests	124
D.1	Simulated density maps based on the IRAS catalogue weighted with the coverage of the PAO	126
D.2	Simulated density maps based on the VCV catalogue weighted with the coverage of the PAO	127

D.3	Simulated density maps based on radio loud galaxies weighted with the coverage of the PAO	128
E.1	Radius dependency of CA I for maps based on catalogues	130
E.2	Radius dependency of CA II for maps based on catalogues	131
E.3	Radius dependency of CA I and II for maps based on radio loud galaxies	132
E.4	Radius dependency of CA I and II for maps based on the ellipse model	133
E.5	Radius dependency of CA I and II for maps based on multipoles . . .	133

Outline

The history of cosmic ray physics started in the beginning of the 20th century. Victor Hess measured an increasing ionization while he was rising with a balloon through the atmosphere whereas the opposite was expected. Another important discovery was made by Pierre Auger and his collaborators as they observed coincident signals on the Jungfraujoeh using detectors which have been separated by several meters. These events led to many follow-up experiments aiming to understand what the nature of the cosmic rays is covering many magnitudes in energy and flux. The largest so far is the Pierre Auger Observatory, located in the Pampa in western Argentina. Its data is the basis of this study. In recent years it provided many major contributions to the field of cosmic ray physics.

In this thesis the distribution of arrival directions of cosmic rays at the highest energies ($\approx 5 \cdot 10^{19} eV$) is examined by means of autocorrelation methods to check if it is compatible with the isotropic expectation. One of the advantages is that no assumptions have to be made about sources and their properties, e.g. distribution and strength. This yields complementary information to the results of the Pierre Auger Observatory regarding the correlation of Ultra High Energy Cosmic Rays (UHECRs) with Active Galactic Nuclei (AGN), which are prominent source candidates. However, further models exist trying to explain this observation. For instance, Centaurus A, a strong radio galaxy, could be responsible for a large fraction of UHECRs. Thus, it is reasonable to detach oneself from these assumptions and concentrate on measured facts, the measured data.

In the first chapter a short introduction to the topic is given, followed by a more general one on cosmic rays. Therein, models of acceleration including possible sources and propagation are explained. The third chapter focuses on the detectors of the Pierre Auger Observatory and the reconstruction of events at the highest energies. In addition, special attention is paid to the monitoring of the High Elevation Auger Telescopes (HEAT), a low energy enhancement of the observatory consisting of three tiltable fluorescence telescopes. With the required background knowledge provided, the next chapter starts with a detailed description of the underlying idea and motivations of the autocorrelation methods. Summarised in chapter 5 are the procedure to generate different Monte Carlo maps and the results obtained by applying the methods to the corresponding samples. For instance, the dependency on the level of background is tested as well as the correlation of the methods among each other. Afterwards, the best method is applied to the dataset comparing the two reconstruction algorithms within the Pierre Auger Observatory. Finally, the findings are discussed and an outlook is given, which concludes this thesis.

1. Introduction

Considering cosmic rays, one of the automatically raised questions is “Where do they come from?”. Do they originate from the solar system, the milky way or other galaxies in our local group? In contrast to photons and neutrinos, which are only minimally deflected during their propagation through space, cosmic rays suffer significant deflections proportional to their charge due to magnetic fields. Hence, the arrival directions of cosmic rays and the positions of their emission in general do not correspond. This makes it hard or even impossible to identify the sources of cosmic rays, especially at lower energies. By increasing the energy the deflections become smaller and a determination of the true origin is more likely. However, due to the steep energy spectrum this leads to a significant reduction of statistic, since the spectrum follows a power law with a spectral index γ of ~ 3 . For instance, if the energy is increased by one order of magnitude the flux is suppressed by a factor of 1000.

To ensure yet a reasonable number of events in the lifetime of an experiment the sensitive area has to be increased drastically. Therefore, all dedicated experiments follow this demand either by using fluorescence telescopes with a large range covering a huge volume of atmosphere and/or surface detectors instrumenting a large area providing a sufficient rate to perform meaningful statistical tests. To identify single sources or hot spots on the sky is still challenging but might become possible in the near future. Of course, this depends strongly on various parameters, e.g. on the acceleration mechanism and interactions occurring during the propagation of the cosmic ray. Instead of focusing on all these variables, one question can be answered without using any of these models:

“Are the arrival directions of UHECRs isotropically distributed?”

Two approaches can be considered to answer this question:

- Correlating Cosmic Rays with astrophysical objects;
- Correlating Cosmic Rays among each other.

Each one has its own assets and drawbacks. While the first one is able to give already a good estimate on the significance with a small number of events, it depends on the chosen catalogue and its potential inaccuracies. Although the contained objects can be proper source candidates, in general the selection is arbitrary in the first place. The second approach based on Autocorrelation methods (considered in this work) is

free of this dependence, but might need more events to be conclusive. Consequently, using these methods removes all unknown parameters from an anisotropy analysis: The only data made use of are the arrival directions and energies, which can be controlled and understood. In many experiments autocorrelation analyses have been and are used, for example within the AGASA and the HiRes experiments. In contrast to them, the Pierre Auger Observatory offers the possibility to perform this kind of analysis with higher statistics than ever before based on the larger instrumented area. Finally, this could add more information to the discussion regarding anisotropy. So far the overall conclusion is unclear, since publications exist reporting the finding of anisotropy as well as the agreement with an isotropic distribution of arrival directions.

2. Coordinate Systems

In astronomy and astroparticle physics many coordinate systems (CS) are in use, each one of them fitting for a special case. In this work four of them occur:

1. local,
2. equatorial,
3. galactic and
4. super-galactic.

It is standard in astronomy when a sky map is shown that the positive x-axis is on the left side. We will, however, use the typical mathematical convention unless stated otherwise.

2.1 Local Coordinates (LCS)

This reference frame corresponds to an earth bound CS centred in the middle of the Pierre Auger Observatory. In general this is referred to as the Horizontal-CS. Since it is rotating, objects on the sky not only need two angles to describe their positions but also an additional time information. Therefore, the coordinates are the azimuth angle ϕ , the zenith angle θ and time t . Per definition $\phi = 0^\circ$ starts in the east going then counter-clockwise, i.e. $\phi = 90^\circ$ corresponds to the north. In addition, θ is zero at the zenith.

The LCS can be considered the detector CS where the measurements of air showers take place. Therefore, it can be used directly to study several detector-related physical observables. However, to access information about their origin it is easier to handle events within a time independent coordinate system.

2.2 Equatorial Coordinates (ECS)

Precisely this one fulfils this requirement and is still geocentric, since the Earth's longitude and latitude are projected onto a sphere with the Earth at its center. Hence, the equator and the poles are arranged in the same way. Coordinates are right ascension α and declination δ , where in contrast to the LCS δ is 90° at the north pole and 0° at the equator.

The advantage of using the Earth's own CS is also one of its drawbacks: Due to

nutations and precessions the positions of objects are time dependent. However, this dependency is rather weak and with time scales much larger than the daily motion of the Earth. To avoid a continuous change of the coordinate system, epochs are introduced. They act as a reference point in time and all celestial objects have fixed positions relating to it. However, after several decades a new epoch has to be introduced due to the increasing distance between true and the referenced position. At the moment the J2000.0 is in use, while in addition B1950.0 can still be found in some catalogues.

2.3 Galactic Coordinates (GCS)

To get rid of this last time dependence, one has to use a stable reference. In the GCS the sun is in the center of the celestial sphere and the equator is chosen to be collinear with the milky way. The longitude l is 0° at the galactic center and the latitude b defined accordingly to declination δ in the ECS. For a detailed information on conversion from one CS to the other see section 2.5.

2.4 Super Galactic Coordinates (SGCS)

Comparable to the GCS, the equator of the SGCS should go through the majority of very massive galaxies in our extragalactic neighbourhood. However, the definition of its orientation is not as straightforward as in the case of the milky way. We follow here the suggestions made in [1] and [2]. In this way, the equator corresponds to the super galactic plane (SGP), which contains the Virgo Cluster and the Great Attractor for instance. In Fig. 2.1 the relative position of the GC and the SGCS is shown including galaxies of the Veron-Cetty and Veron catalogue [3]. While close to the equator of the SGCS several galaxies can be found, the opposite is true for the equator of the GCS (red dashed line). This absence can be explained by the fact that the Milky Way obscures the sight on galaxies lying behind it.

For an anisotropy analysis which depends on the reference system this is considered a reasonable choice. Since, close to the SGP many galaxies can be found and some of them are possible source candidates for UHECRs. This could lead to an overabundance of cosmic rays in the vicinity of the equator of the SGCS.

2.5 Transformation

For converting coordinates from one system into another it is necessary to have at least the positions of two reference points in both cases. Here, we present those we used to rotate the coordinate systems.

	SGCS (L,B)	GCS (l,b)	ECS (α,δ)
1. Coordinate	($0^\circ,0^\circ$)	($137.37^\circ,0^\circ$)	($283.5^\circ,59.5^\circ$)
2. Coordinate	($0^\circ,90^\circ$)	($47.37^\circ,6.32^\circ$)	($42.3^\circ,15.7^\circ$)

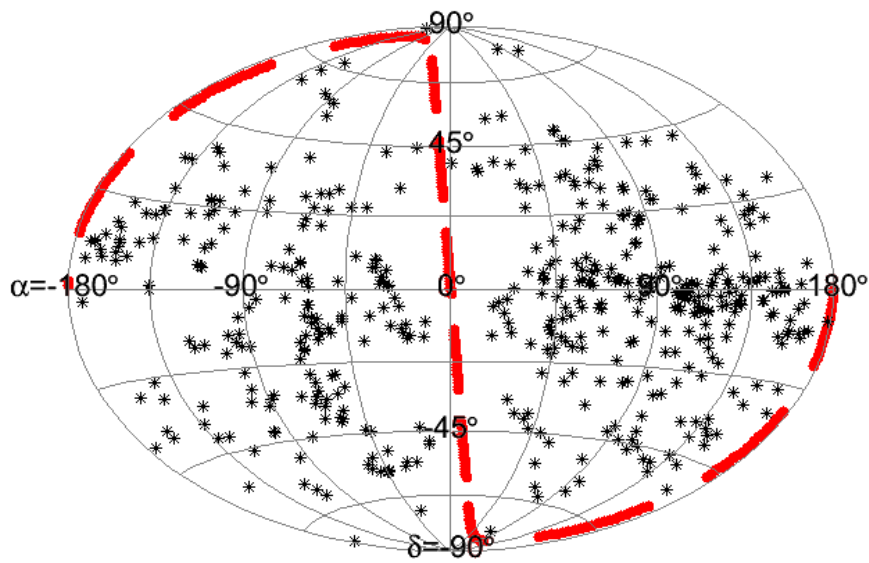


Figure 2.1: As an example galaxies contained in the Veron-Cetty and Veron are shown in the SGCS with a redshift smaller than 0.02. Several galaxies are found in the vicinity of the equator. The red dashed line represents the equator of the GCS transformed into the SGCS.

3. Cosmic Rays

In the beginning of the 20th century particle physics was mainly performed by studying cosmic rays, which are defined as charged and stable particles (ionized nuclei). First discovered by Victor Hess in 1912 [4], they offered the opportunity to investigate elementary particle physics with energies exceeding the MeV regime, because man-made particle accelerators delivering these energies were unavailable until the 1950s. In this way for example new particles were discovered like the positron in 1932 [5] and the muon in 1936 [6], both by C. D. Anderson and the latter in collaboration with S. H. Neddermeyer.

Over the past 100 years the measured energy range has been expanded by many experiments, airborne as well as ground-based, from a few GeV up to 10^{11} GeV covering a differential flux $d\Phi/dE$ over 30 orders of magnitude (see Fig. 3.1). In general, $d\Phi/dE$ follows a power law $E^{-\gamma}$ with a varying spectral index $\gamma \in [2.7, 3.2]$. Over the complete range a few features occur where γ changes. These features are of special interest because a change in the underlying physical processes accelerating and affecting cosmic rays can be assumed. The best studied kink is called the knee around $5 \cdot 10^{15}$ eV due to the still large flux. Further discussion on this subject can be found in chapters 3.2 and 3.3.

As one can see from Fig. 3.1, the flux is dominated by the lowest energies, which are mainly protons with a flux of ca. $1000 \text{ m}^{-2}\text{sr}^{-1}\text{s}^{-1}$ above the atmosphere.

Regarding the composition of cosmic rays, the same elements have been found in cosmic rays as in our solar system with nearly the same abundance.

Direct measurements are only possible at lower energies ($E \lesssim 10^5$ GeV) with balloons and satellites where the flux is sufficiently high. For higher energies one uses extensive air showers (EAS) induced by cosmic rays to reconstruct energy, arrival direction and composition: A primary particle, a proton for example, interacts with the nuclei of nitrogen and/or oxygen in the Earth's atmosphere producing secondary particles which in turn produce further particles until the energy per particle is below the threshold. The concept of EAS is the basis for indirect measurements explained in more detail in chapter 3.4. By using ground based detectors covering a huge area, the small flux can be partly compensated, which is necessary especially at the highest energies, $E \approx 10^{20}$ eV. Then, the rate drops to a value of $1 \text{ km}^{-2}\text{century}^{-1}$. Estimation of the correct values during the reconstruction relies on the understanding of the processes happening in the atmosphere. Therefore, precision measurements of accelerator experiments of cross-sections for example are essential to predict the correct behaviour. The results are integrated into hadronic interaction programs (EPOS [7], QGSJETII [8]). However, these measurements are limited to the energies of the order of 1 TeV in the center of mass system until now and focus on

proton-proton collisions whereupon the targets as well as some UHECRs are nuclei. Therefore, the obtained results must be extrapolated to higher energies and to other processes leading to larger uncertainties. Furthermore, an interaction model alone is not sufficient to describe an EAS. Hence, a modelling of the complete reaction chain through simulation programs (CORSIKA [9], CONEX [10], Aires [11]) is needed which incorporates the interaction models.

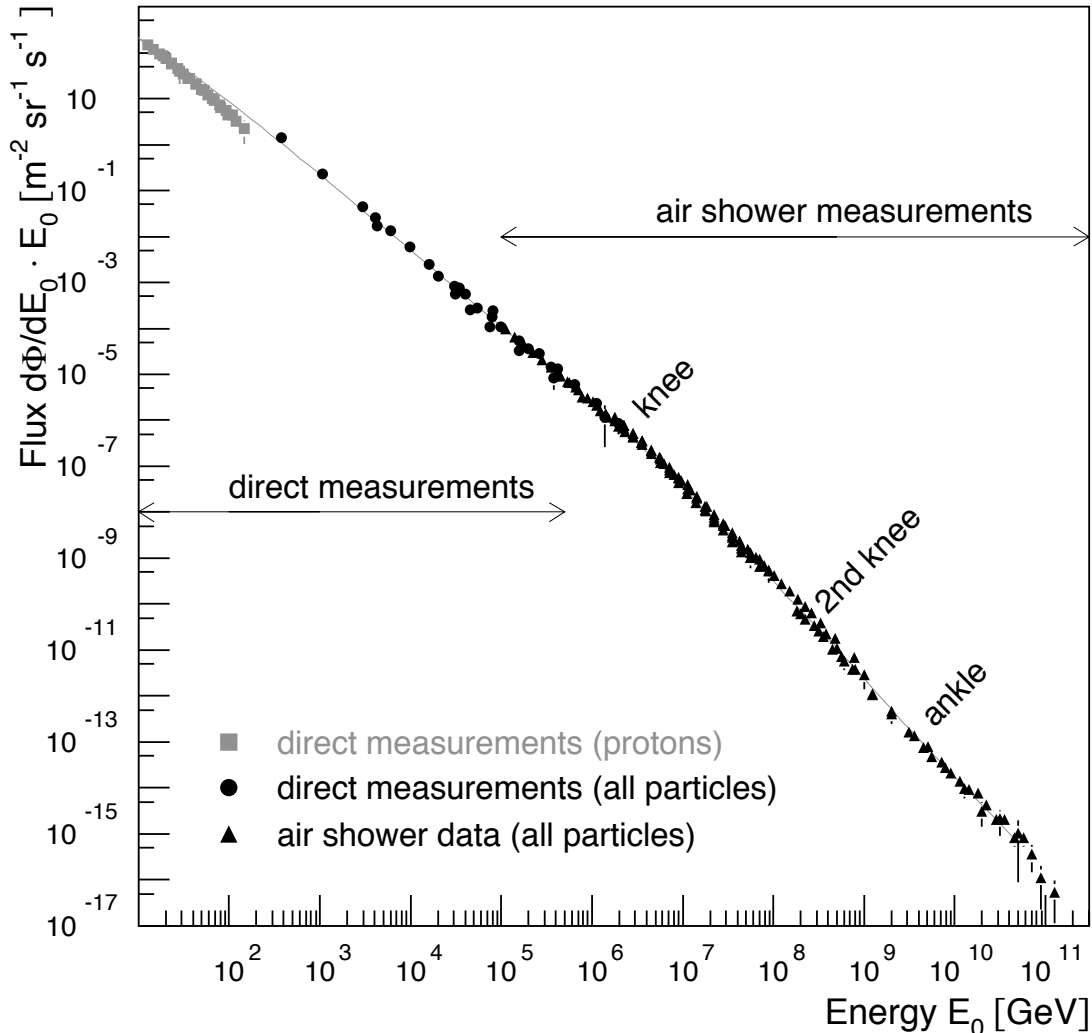


Figure 3.1: Flux of cosmic rays for all nuclei separated by direct and indirect measurements [12].

This applies in particular to the measurement of the composition, because the nature of the primary particle can be estimated by studying the shower development by means of the shower maximum X_{max} where the number of particles within the shower is maximal. The change in energy of X_{max} is called the elongation rate, dX_{max}/dE and is used to study the development of the composition with energy. Thus, where a classification is possible at lower energies in elemental groups, at the

highest a distinction in light and heavy is already more complicated. Recent results of the Pierre Auger Observatory (PAO) regarding the composition are shown in Fig. 3.2. At energies around 10^{18} eV the error for the simulations is rather small and the cosmic rays seem to be light ($Z \approx 1$). For higher energies the data tends to a heavier composition. However, to draw a conclusion is difficult, which is in part caused by the increased uncertainties of the models. The RMS of the X_{max} distribution on the other hand (right panel of Fig. 3.2) hints towards an interpretation of the data as heavy, since the systematic error on the predictions is smaller and the data points are showing a clearer trend. On the other hand, the HiRes experiment [13] reported in 2010 their final results including an agreement of their data with a pure proton composition. Efforts are under way to solve these discrepancies.

Although, the knowledge about cosmic rays has increased over the last years many questions are still unresolved. Especially, the origin of cosmic rays at the highest energies is still unknown and with it entwined the question concerning the composition. Protons would make astronomy with UHECRs possible, while in the case of iron the deflection due to magnetic fields would be 26 times larger leading to a more isotropic distribution of the arrival directions.

Before discussing possible source scenarios it is necessary to explain a mechanism which could accelerate cosmic rays to the highest energies. In this way, the number of source candidates can be narrowed down. This will be followed by sections on propagation and extensive air showers.

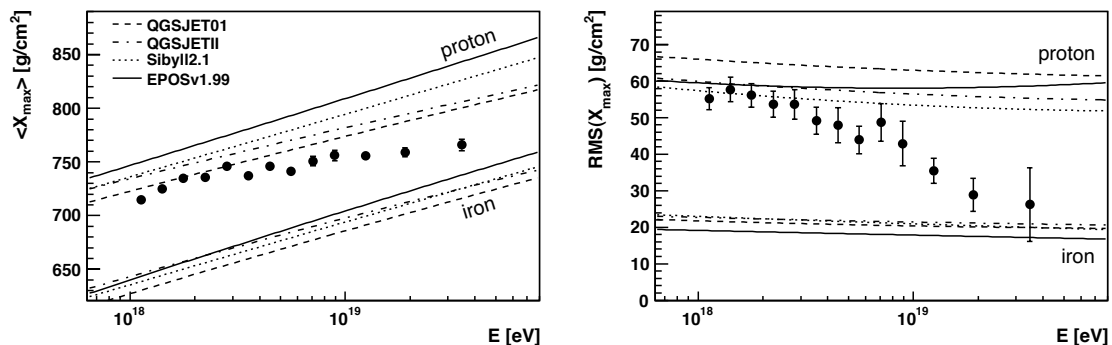


Figure 3.2: Composition Measurements of cosmic rays compared to the predictions (lines) made by shower simulations. *Left Panel:* X_{max} measurement of PAO indicating a light composition (protons) at lower and a mixed one at higher energies. *Right Panel:* RMS of X_{max} measurement of PAO with a clear hint towards iron at the highest energies. Both taken from [14].

3.1 Fermi Acceleration

In 1949 Enrico Fermi [15] published a first approach to explain the existence of cosmic rays with very high energies based on stochastic acceleration. This theory has been refined in the following decades and is still under investigation. New developments for example gave rise to a more efficient mechanism accelerating cosmic rays.

In both cases that are presented in the following, a change in energy is based on a

repeatedly encounter with a magnetic environment. Each one leads to an increase in energy of the amount ϵ on average, i.e. after n deflections the relative gain is $(1 + \epsilon)^n$. To calculate the number of steps necessary to reach a designated maximal energy one uses

$$n = \frac{\ln(E_n/E_0)}{\ln(1 + \epsilon)}.$$

Another important parameter is the observed power law for the flux. Therefore, an acceleration mechanism claiming to be responsible for the existence of UHECRs has to produce such a behaviour.

In this case, the mechanism can provide a power law in the end with the spectral index being connected to the gain in energy per encounter. This can be shown by considering the number of particles staying in a region where several successive accelerations can take place. For one particle the chance to escape this area is p_{esc} per encounter with the magnetic cloud. From that we can conclude that after n interactions the number of particles is reduced by a factor of $(1 - p_{esc})^n$. By summing over n until infinity the number of particles achieving a higher energy than E_n is obtained:

$$N(> E_n) = N_0 \sum_{m=n}^{\infty} (1 - p_{esc})^m.$$

This is a geometric series and can be simplified towards

$$\begin{aligned} N(> E_n) &= N_0 \left(\frac{1}{1 - (1 - p_{esc})} - \frac{1 - (1 - p_{esc})^n}{1 - (1 - p_{esc})} \right) \\ &= \frac{N_0}{p_{esc}} \cdot \exp\left(\frac{\ln(E_n/E_0) \cdot \ln(1 - p_{esc})}{\ln(1 + \epsilon)}\right) \\ &= \frac{N_0}{p_{esc}} \cdot \left(\frac{E_n}{E_0}\right)^{-\gamma} \end{aligned}$$

with

$$\gamma = \frac{\ln(1/(1 - p_{esc}))}{\ln(1 + \epsilon)} \approx \frac{p_{esc}}{\epsilon}$$

whereupon the last equality is only valid for small values of p_{esc} and ϵ .

3.1.1 Second Order Fermi Acceleration

This represents the original model by Fermi [15] and considers the interaction of interstellar magnetic clouds and relativistic cosmic rays. In this simple example (adapted from Stanev [16]) cloud and cosmic ray are moving towards each other colliding head-on. The particle is deflected many times inside the cloud and escapes in the same direction from which it came. By transforming the incoming cosmic ray with energy E_s from the laboratory into the reference frame of the cloud (indicated

by a superscript cl) and afterwards the other way round, the change in energy E_f can be calculated. In this simple example, it is:

$$\begin{aligned} E_s^{cl} &= \gamma_{cl} E_s (1 + \beta^{cl}) \\ E_f &= \gamma_{cl} E_s^{cl} (1 + \beta^{cl}) \end{aligned}$$

with the corresponding relativistic γ - and β -factors of the cloud. Hence, one can calculate the increase of energy ϵ which is depending on the cloud velocity v_{cl} :

$$\epsilon = \gamma_{cl}^2 (1 + \beta_{cl})^2 - 1,$$

with $\beta_{cl} = |v_{cl}|/c$, $\gamma_{cl} = (1 - \beta_{cl}^2)^{-\frac{1}{2}}$ and c , the speed of light.

In general, the direction of escape is isotropically distributed due to many deflection inside the cloud giving in the end an average gain in energy of

$$\epsilon \approx \frac{4}{3} \beta_{cl}^2.$$

From that we can conclude that the spectral index is determined among other things by the squared velocity of the cloud in units of the speed of light. Hence, the time to accelerate particles depends on the number of encounters with a magnetic cloud and its velocity. Due to the variability of v_{cl} , a prediction for γ is impossible.

3.1.2 First Order Fermi Acceleration

This kind of acceleration takes place in objects like Supernovae Remnants (SNR), where a shock front hits the surrounding interstellar medium. Cosmic rays within this matter are isotropic in their flying direction. In the rest frame of the shock some cosmic rays cross the shock-front head-on with the velocity v_{shock} and will be accelerated according to the principle of Fermi of second order where a cosmic ray interacts with a magnetic cloud (see the previous section). After the cosmic ray is isotropized in the shocked medium it can cross the shock front for the second time into the unshocked medium where it is accelerated again according to Fermi of second order. This process can recur many times, whereupon the average gain in energy is:

$$\epsilon \approx \frac{4}{3} \beta_{SM} = \beta_S. \quad (3.1)$$

It must be noted that this is the β_{SM} -factor of the shocked medium, which usually is 3/4 times the shock velocity v_S .

Due to its similarities to the original concept of Fermi, this mechanism is called first order Fermi acceleration. A description of this acceleration process was published independently by several people in the 1970's, e.g. by Krymskii [17] as well as Blandford and Ostriker [18].

As one can see from equation (3.1) the first order Fermi acceleration is more effective, since the average shock velocity within a SNR is much higher than the one of typical interstellar molecular cloud and in addition the gain in energy is linear in

the β_S -factor of the shock.

Strong shock fronts produce in a natural way $\gamma \approx 2$ [18] at the source. During the propagation of cosmic rays through a galaxy, some of them escape the confinement of the galaxy. Solving the corresponding diffusion equation, this leads to an increase of the spectral index of 0.6 (see [16]). Further effects, which are discussed below, change again the spectral, especially at the highest energies.

3.2 The Origin of Cosmic Rays up to the knee

Due to the magnetic fields in our galaxy cosmic rays are deflected many times so that the complete source information is lost. The diffusive flux however can still give hints towards a solution of the question about the origin in this case. Particularly, changes in the spectral index are interesting and could point to a change in the underlying processes. A good example is the kink around $5 \cdot 10^5$ GeV called the knee, where the spectral index γ changes from 2.7 to 3.2. Prominent candidates for this energy range are Supernovae in our own galaxy. Particles gain energy by means of Fermi-acceleration of first order while passing through the shock-front many times back and forth.

Two scenarios are discussed in the literature explaining the steepening of the flux. The first one argues that the accelerators run out of power leading to a cut off for each particle species with a maximum energy of $E \approx Z \cdot 10^{14}$ eV proportional to their charge. However, depending on the theory this value can be enlarged by up to three orders of magnitude.

In the second scenario the galactic magnetic field B is too small to confine cosmic rays of energy E and charge Z within so that the Larmor radius

$$r_L = 1.08 \text{pc} \frac{E/\text{PeV}}{Z \cdot B/\mu\text{G}} \quad (3.2)$$

exceeds the size of the disc of the milky way and the cosmic rays can escape the boundaries of our galaxy. Since the corresponding radius r_L of a cosmic ray depends on its charge, nuclei of higher atomic order can be kept inside the galaxy up to higher energies than it is possible for protons. By measuring the composition and the corresponding fluxes in the transition region this hypothesis can be verified.

At the moment one assumes that in nature a mixture of both theories is realised.

3.3 The Origin of UHECRs

Going to even higher energies, further features become apparent. To make them clearly visible, the differential flux is multiplied by $E^{2.7}$ (see Fig. 3.3). Since galactic sources are assumed to die out at lower energies as discussed in the previous section, extragalactic objects are held responsible for the flattening of the spectrum around $E = 3 \cdot 10^{18}$ eV, called the ankle (cf. Fig. 3.3). In contrast to the accelerators within a galaxy like supernovae, centres of galaxies can offer a good environment for particle acceleration; the focus lies on the active galactic nuclei (AGN). In addition,

more exotic objects are included in this category, e.g. fast rotating neutrons stars, colliding galaxies and Gamma Ray Bursts (GRB) which can also be located inside our own galaxy. In contrast to lower energies, further models exist favouring among other theories very heavy particles based on the supersymmetric extension of the standard model of elementary particle physics (abbreviated SUSY) or topological defects produced in the early universe. Through annihilation or decay UHECRs are then produced. In general, these two approaches are distinguished into:

1. bottom-up (based on extra- and galactic objects, e.g. AGN and GRB) accelerating cosmic rays from moderate up to the highest energies and
2. top-down models (based on particles, e.g. decaying super heavy dark matter), energy provided by the mass of the particle.

Both are discussed in more detail below. Many important experiments have contributed so far to this field of UHECR physics. To give a short incomplete list: Haverah Park (1967-1987), AGASA, HiRes (1997-2006) and of course the Pierre Auger Observatory (running since 2004). Most results presented in this chapter are based on this last experiment, because it is the largest so far delivering current measurements with higher statistics than any experiment before.

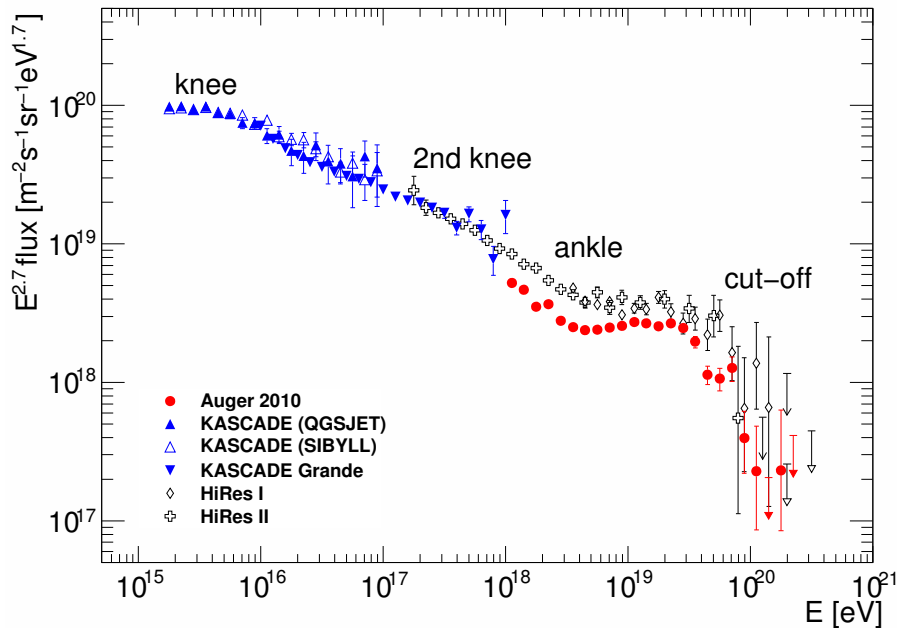


Figure 3.3: Flux of cosmic rays multiplied with $E^{2.7}$ to enhance the features significantly using KASCADE [19], KASCADE Grande [20]. HiRes [21] and PAO data [22]. (provided by M. Grigat and C. Meurer)

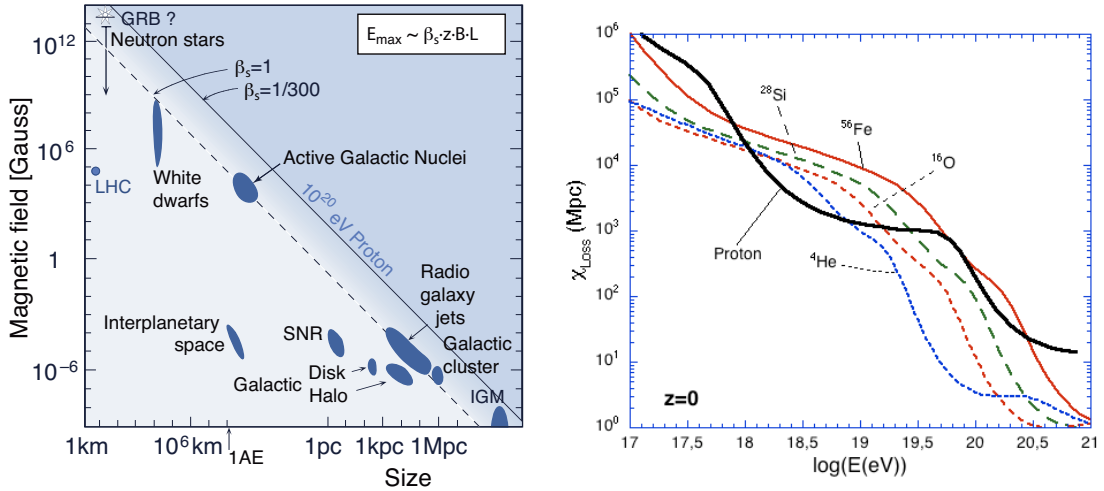


Figure 3.4: *Left Panel:* Shown is the Hillas Plot which connects the size of possible source candidates and the required local B-field to accelerate UHECRs according to equation (3.3). Diagonals represent different shock velocities (adapted from [12], original [23]). *Right Panel:* Expected loss length for some nuclei. For protons the first dip (e^+e^- -production) is clearly visible followed by the GZK cut-off. (taken from [24])

3.3.1 Bottom-Up Models

Within this category all macroscopic objects are combined being possible acceleration sites for UHECR. All these objects have in common that they fulfil the Hillas relation [23] based on Fermi acceleration of first order:

$$E_{max} \simeq 10^{18} \text{eV} Z \beta_S \left(\frac{R}{\text{kpc}} \right) \left(\frac{B}{\mu\text{G}} \right), \quad (3.3)$$

where Z is the charge of the particle, $\beta_S = \frac{v_S}{c}$ the shock velocity, R the size of the accelerating region and B the magnetic field. Applied to known objects possible source candidates are obtained. A selection can be found in Fig. 3.4 including Active Galactic Nuclei (AGN) and Gamma Ray Bursts (GRBs).

3.3.1.1 Active Galactic Nuclei

Compared to our own, other galaxies' most inner part is more active shining brighter at different wavelengths (e.g. radio, infrared, gamma-ray). A possible reason for this is the presence of an accretion disc surrounding the central super massive black hole. This matter consisting of gas and dust is drawn into the black hole increasing pressure and friction in the torus, whereupon heat and radiation are produced. In some cases, further energy is released through two jets erupting back to back perpendicular to the accretion disc along the rotation axis, whereas the jets are often asymmetric in their shape and size. Hence, AGN have also been observed with only one jet. The mechanism behind this phenomenon is still subject of research.

With the help of Fig. 3.5 the different parts of an AGN can be explained in more

detail. AGN are divided into two major subclasses displayed here by a diagonal line, radio loud and radio quiet galaxies. The significant difference is the presence of the already mentioned jet in the first case strongly emitting in the radio band. When looking directly into the jet the object is called a *Blazar*. In the radio quiet mode the galaxy would be called a *Quasar*.

Further criteria of classification are based on the abundance of narrow and/or broad optical lines in the measured photon spectra. The reason for these two types is cold matter rotating around the black hole, the closer it is the faster it spins and thus the larger is the difference of the matter velocities. Therefore, the Doppler shifts are varying more and hence the emission lines are broader. In principal, both subtypes of AGN, Seyfert galaxies of type 1 and 2, are the same, only the viewing angle is different. While in the case of Seyfert type 2 just the narrow lines can be seen, for type 1 both are visible. By now, the sharp partition is relaxed and a relative nomenclature is in use, i.e. type 1 and 2 are limits and numbers in between are more common.

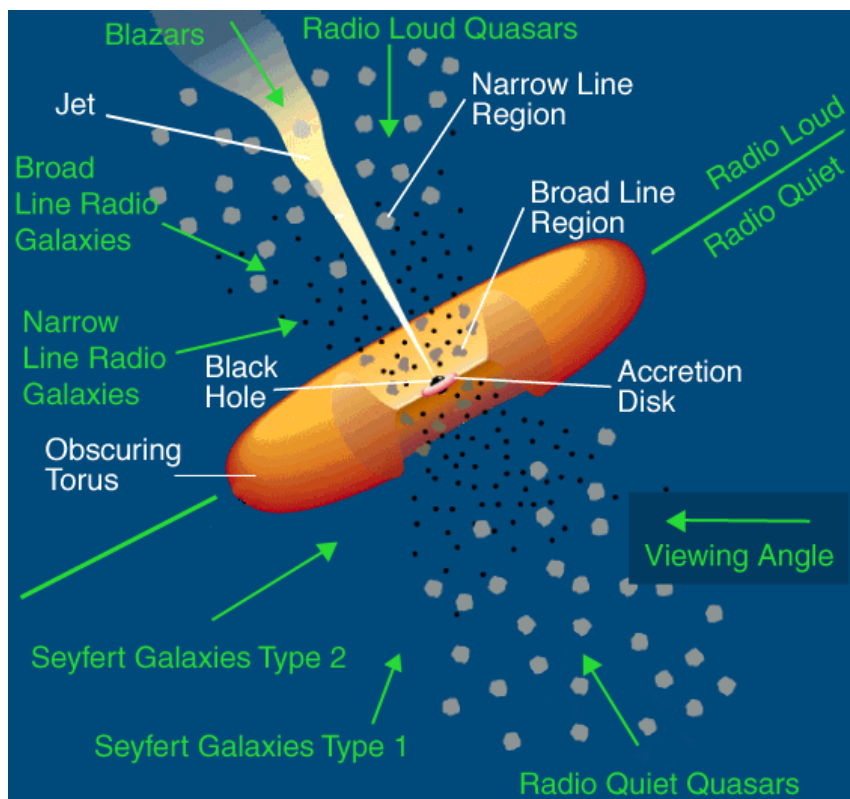


Figure 3.5: Scheme of Active Galactic Nuclei. A massive black hole in the center surrounded by an accretion disc and a gas and dust torus. Interstellar matter rotates around the black hole producing optical narrow as well as broad lines depending on its velocity. In the case of a radio loud galaxy a jet is emitted along the rotation axis. (provided by NASA)

Much attention is given to these objects caused by the publication of the PAO regarding the correlation of their 69 highest UHECRs with AGN of the Veron-Cetty

and Veron (VCV) [3] catalogue. After two years of data taking a scan was performed in energy E , opening angle ψ around the galaxies and redshift z to find the best values for a prescription regarding a correlation. They are as follows:

$$\begin{aligned} E &\geq 56 \text{ EeV} \\ \psi &= 3.1^\circ \\ z &\leq 0.018 (\approx 75 \text{ Mpc}) \end{aligned}$$

With this data set a relative correlation of $\rho = 57\%$ was obtained, whereas only 21% were expected for events coming from an isotropic distribution (H_0 -hypothesis). The probability whether the arrival directions of the detected UHECRs are isotropically distributed is calculated by means of the cumulative binomial distribution. In this context, the errors of first and second kind are set to 1% and 5% respectively. In this way, the prescription ends after 34 events are detected surviving the lower energy cut. Assuming that the values found with the scan are correct, then 15 events should be closer to an AGN than 3.1° . Due to the small flux of UHECRs at the highest energies, it was decided to test after the detection of a new event whether the prescription is already fulfilled. To account for the repeated testing, the error of first kind was reduced by a factor of four.

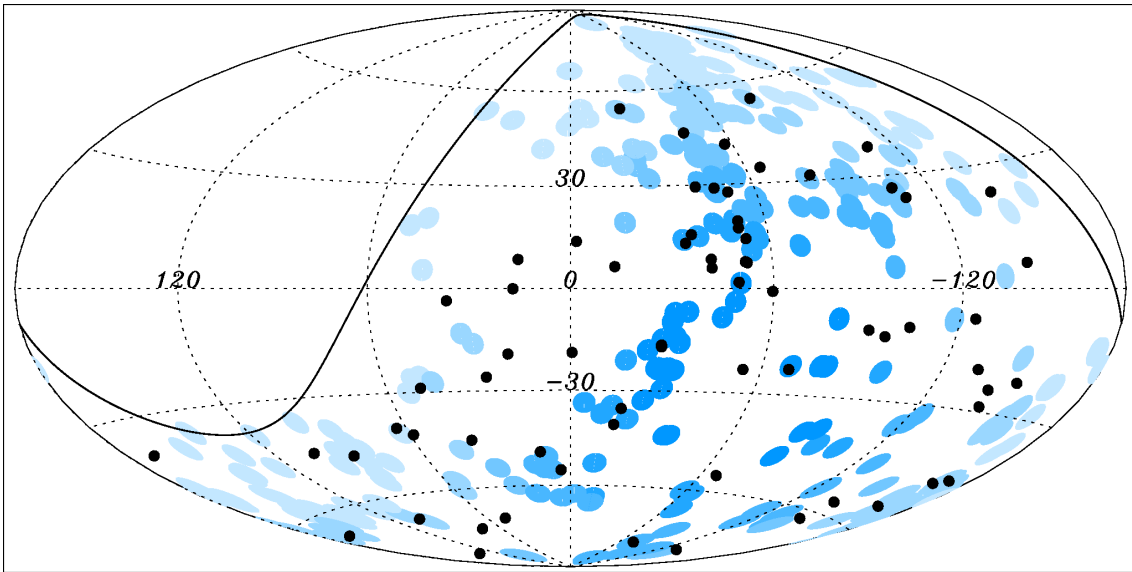


Figure 3.6: Shown are the 69 highest energetic PAO events in black (galactic coordinates, Hammer projection). In addition all VCV AGN with $z \leq 0.018$ are shown as blue circles their shade proportional to the exposure. [25]

Already in 2007, the prescription passed (published in [26]), with $\rho = 70\%$ giving a chance probability P_{chance} of 0.02% that the events are isotropic. Shortly after the publication, the amount of correlating events dropped to an overall value of ca. 40% reported in 2010 [25], which is still a factor two higher than expected by H_0 -hypothesis (isotropy). The chance probability increased ($P_{chance} = 0.3\%$). In Fig. 3.6 all events of the PAO above an energy of 55 EeV are shown as

Table 3.1: Summary of correlations within 3.1° between cosmic rays with $E \geq 55$ EeV and AGNs in the VCV catalog with redshift $z \leq 0.018$. N is the number of CRs measured. k is the number of correlating arrival directions. k_{iso} is the number of correlations expected by chance if the flux were isotropic. P is the cumulative binomial probability to detect k or more correlations from an isotropic distribution. Probabilities are not shown for data sets which include period I because parameters were selected to optimise the correlation in that period.[25]

Period	Dates	Exposure km ² sr y	N	k	k_{iso}	P
I	1 Jan 2004 - 26 May 2006	4390	14	8	2.9	–
II	27 May 2006 - 31 Aug 2007	4500	13	9	2.7	2×10^{-4}
III	1 Sept 2007 - 31 Dec 2009	11480	42	12	8.8	0.15
Total	1 Jan 2004 - 31 Dec 2009	20370	69	29	14.5	–
II+III	27 May 2006 - 31 Dec 2009	15980	55	21	11.6	$\mathbf{3 \times 10^{-3}}$

well as all AGN from the VCV with a $z \leq 0.017$. They are displayed as blue circles with a radius of ψ and their shade proportional to the exposure in that region. An overview about the single periods including probability and all other necessary variables is given in Tab. 3.1.

It must be noted that despite the correlation the true sources stay unidentified, since the statistic is insufficient to relate the events to certain single objects. Crosschecks performed with further catalogues, show a similar correlation leading to the conclusion that the sources follow the matter distribution in our cosmological neighbourhood of 100-200 Mpc. Particularly interesting is the accumulation of events close to a part of the super-galactic plane where lots of AGN can be found in Fig. 3.6.

However, these results are disputed by the observations of the HiRes-collaboration [27][28], because in their data set no such correlation can be found. Two out of their 13 events correlate with VCV galaxies using the same cuts as PAO giving an agreement with the expected isotropic flux of 85%. One explanation for this might be the different fields of view, since HiRes, in contrast to PAO, is located at the northern hemisphere.

Radio Loud Galaxies

The existence of jets and thus of shock fronts makes radio loud galaxies, a subclass of AGN, prominent source candidates. It has been proposed that such galaxies can be responsible for the acceleration of UHECRs. The three with the strongest flux at 5 GHz at Earth within a distance of 100 Mpc according to P. Biermann [29] are Centaurus A (Cen A), Virgo A (Vir A) and Fornax A (For A) (see Tab. 3.2). Especially interesting is Cen A, because it is the closest of those three and the largest object seen in the radio sky for an earthbound observer. Responsible for this are the two large radio lobes expanding far into the intergalactic medium. While the moon has a diameter of 0.5° , Cen A spans a range of more than 10° .

Proposed by theorists (see e.g. [30], [31]), this object may be responsible for a significant amount of all UHECR measured by the PAO. For the HiRes experiment on

Table 3.2: The strongest radio-loud galaxies in the vicinity of the milky way. Given are name, distance d and the equatorial coordinates α and δ

Name	d [Mpc]	α [°]	δ [°]
Cen A (NGC 5128)	ca. 3.5	201.37	-43.02
Vir A (NGC 4486)	16.4	187.70	12.39
For A (NGC 1316)	ca. 19.0	50.67	-37.21

the other hand Cen A is close to their exposure boundary reducing the probability to see events from that direction.

Indeed, the PAO observes many events in its vicinity. However, Cen A is part of the super galactic plane and several other galaxies are found in the same direction with a distance smaller than 200 Mpc. Because of that, a clear identification as a source is challenging. Nevertheless, the PAO has analysed a posteriori in [25] the cumulative angular distance distribution of the events to Cen A where ψ_{CenA} is the maximal allowed angular distance. The PAO has found a small excess at roughly 20° (cf. Fig. 3.7) compared to the isotropic expectation. This number seems to confirm the acceleration rather in the jets than in the nucleus of the galaxy. Further statistics will shed more light on this question.

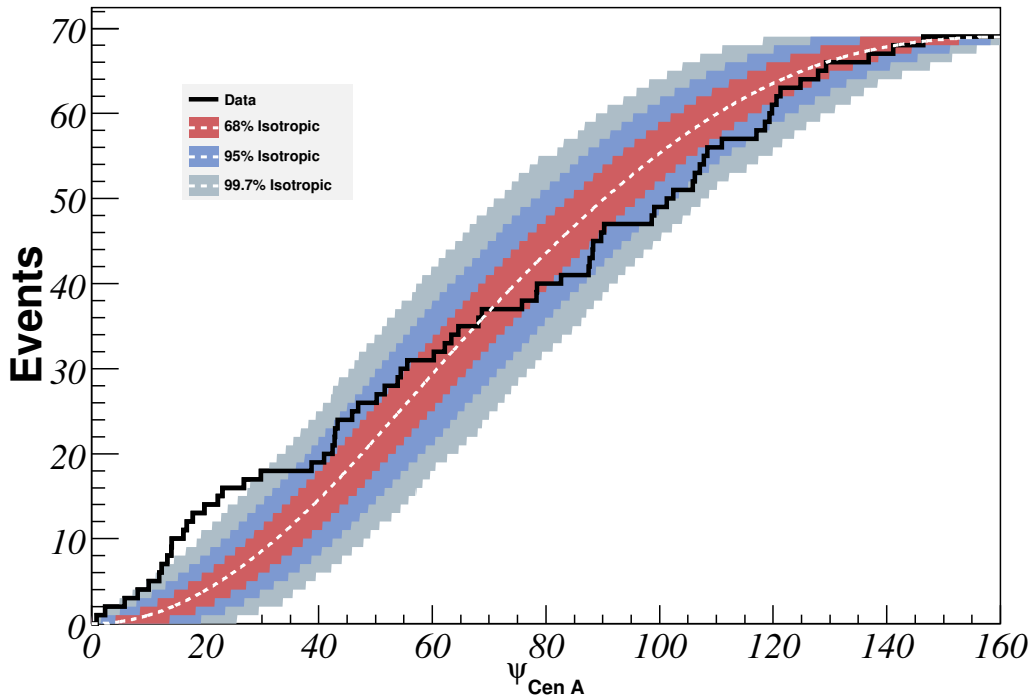


Figure 3.7: Cumulative Distance Distribution for the highest energy events around Cen A including the 1σ , 2σ and 3σ confidence bands for an expected isotropic flux [25]. ψ_{CenA} is the maximal allowed angular distance between Cen A and UHECRs detected by the PAO.

3.3.1.2 Gamma Ray Bursts

In contrast to the long lifetime of an AGN, the lifetime τ of a GRB is only several milliseconds up to hundreds of seconds. These GRB are classified into two groups, long ($\tau > 2\text{s}$) and short ($\tau < 2\text{s}$) each followed by a less intense afterglow. While the first one is believed to be the result of a collapsing massive star forming a black hole, the second one is most likely caused by the merging of two neutron stars.

During these processes a relativistic shockwave is released into two back to back jets hitting the surrounding interstellar matter. Mechanisms like the fireball [32],[33] or the cannonball [34] model describe the occurring reactions in the shockwaves accelerating charged particles, e.g. protons up to $E = 10^{20}$ eV. However, it can be argued that distant GRB cannot contribute to the flux above 10^{19} eV and those which are close-by are too rare to be responsible for the measured flux of UHECRs.

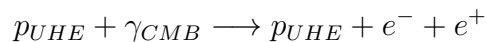
3.3.2 Top-Down Models

In contrast to the concept before, Top-Down models propose a microscopic origin of UHECRs. Many approaches are made starting from topological defects to super heavy dark matter. All these models have in common that at the highest energies photons contribute a significant amount to the flux of UHE particles. Therefore, by searching the data for photon candidates these models can be directly tested.

In 2009, the PAO has published the updated results of their search for photons pushing the upper limits of the photon fraction below the expectation for lots of those models (cf. Fig. 3.8). Clearly disfavoured are the super heavy dark matter models and in addition those based on topological defects. Only the Z-Burst scenario and photons as a result of the GZK-effect (see chapter 3.3.3.1) are still consistent with the data. A Z-burst is understood to be caused by UHE neutrinos interacting with relic anti-neutrinos producing resonantly a Z-boson which in turn decay indirectly into photons (of the order of 10 %).

3.3.3 Propagation

During the propagation of UHECRs through space they are deflected by many effects mainly caused by interactions with the cosmic microwave background (CMB). The most famous one is the GZK-effect at the highest energies leading to an immense energy loss for the UHECRs. Another process is already active for energies of 10^{18} eV namely pair production of UHECRs with the CMB photons.



The reason why it is not so prominent is that the protons lose just about 0.1 % of their energy during this process with a loss length χ_{Loss} of the order of 1 Gpc. χ_{Loss} corresponds to the average propagation distance where the particle has only $1/e$ of its initial energy left.

Another effect which causes only minimal energy loss is the adiabatic expansion of the universe. It can be neglected when only small distances up to 200 Mpc are considered.

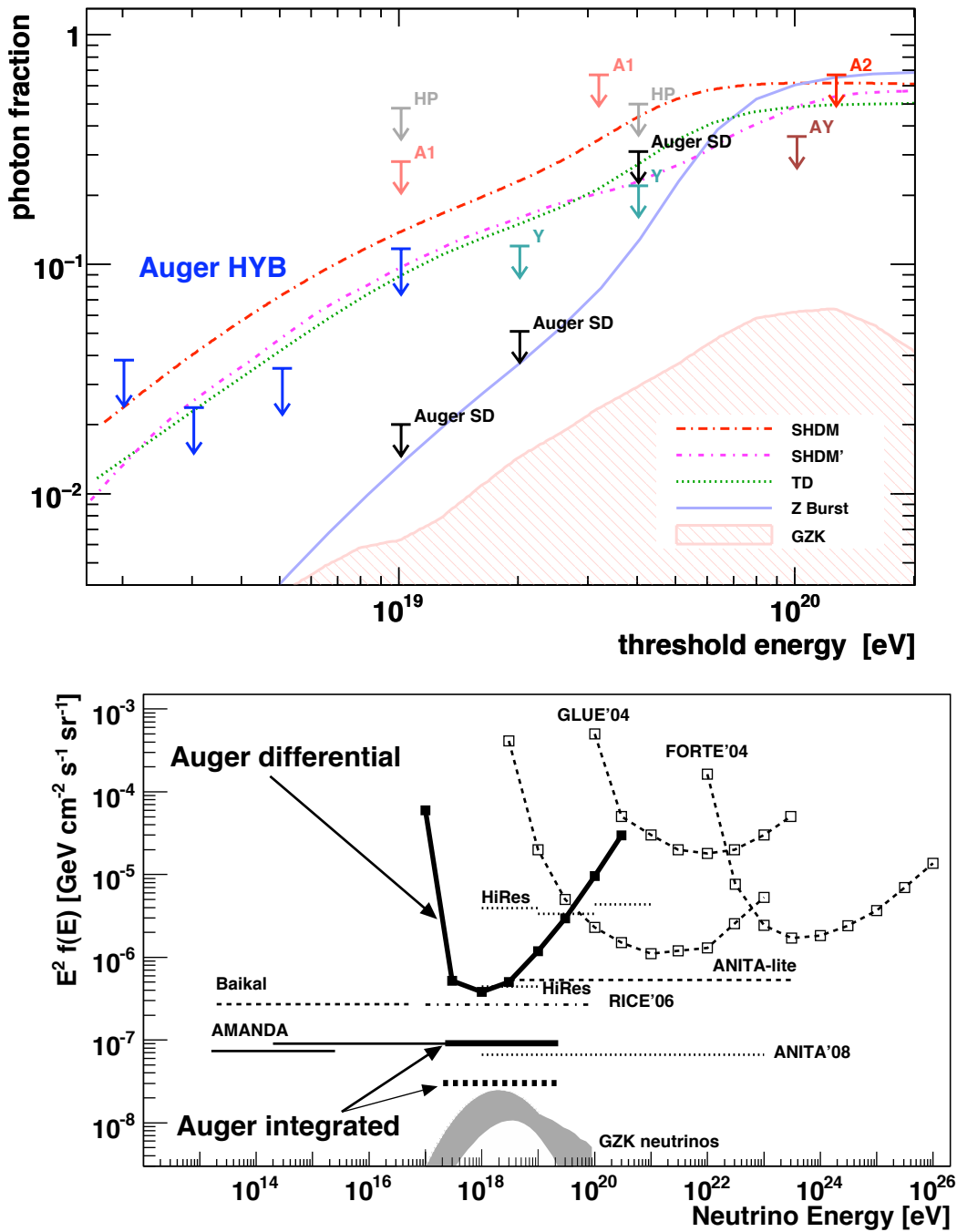
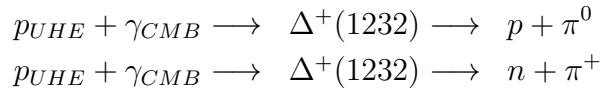


Figure 3.8: *Top Panel:* Current Photon limits of the PAO using data of the surface detector (SD) and the fluorescence detector with one surface detector station (HYB) excluding super heavy dark matter and topological defect models. Only the Z-Burst model and the predictions of the GZK effect (see chapter 3.3.3.1) are still allowed. In addition, results of other experiments are shown, e.g. AGASA (A1, A2), AGASA-Yakutsk (AY), Yakutsk (Y), Haverah Park (HP). (See [35] and references therein.) *Bottom Panel:* Current τ -neutrino limits of PAO using only SD data assuming a ratio of 1:1:1 among the three families. Shown are the differential upper limit on the flux as well as the integrated one over energy, whereupon the dashed line in the integrated case represents the most optimistic and the straight line the most pessimistic case for systematic uncertainties. In addition, results of other mainly neutrino-dedicated experiments are shown. (See [36] and references therein.)

3.3.3.1 GZK-Effect

In 1966 Greisen [37] as well as Zatsepin and Kuzmin [38] described a reaction regarding cosmic rays and photons of the Cosmic Microwave Background (CMB). Protons can produce a Δ -resonance when exceeding a threshold energy of $5 \cdot 10^{19}$ eV.



In this way, they lose up to 20% of their energy to the decay products of the pions namely neutrinos, photons and charged leptons, whereas the neutral particles are subject of further research. At the moment, current experiments are getting closer to the predicted fluxes for those particles (cf. Fig. 3.8). The corresponding loss length for the process mentioned above is of the order of 20 Mpc.

Furthermore, nuclei suffer from similar reactions like pair- and photoproduction decreasing their energy. In addition, they can sustain photodisintegration caused by CMB photons. An estimation of the average loss length for some nuclei is shown in Fig. 3.4. As one can see, if an experiment measures cosmic rays with energies exceeding 10^{20} eV then their origin lies most likely within a distance of 100 Mpc. However, processes have been proposed which can suppress the GZK-effect and shift this cut-off to even higher energies including the Z-Burst scenario mentioned before as well as Lorentz violation (for further reading regarding these and other effects, see [39]). Current measurements from HiRes [21] and the PAO confirm the prediction of the GZK (cf. Fig. 3.3). The flux is suppressed for energies above $4 \cdot 10^{19}$ eV at the level of 20σ [22]. AGASA on the other hand published their energy spectrum of cosmic rays which continues beyond $5 \cdot 10^{19}$ eV [40] showing no steepening of the cosmic ray flux. One reason for these differing results might be the uncertainty regarding the energy scale which depends on the experiment. It has been shown that by adjusting the energy scales of these three experiments with factors of roughly 20 % an agreement can be achieved (e.g. Fig. 12 in [12]).

Regarding the photons and neutrinos produced during the GZK process, the first type interacts with the radio background and the second one remains mainly unaffected. For both cases, predications are made regarding the resulting fluxes but no candidate has been found so far in the data of neither the PAO nor other experiments. The current limits can be seen in Fig. 3.8. However, if finally a positive signal from photons or neutrinos is detected, this would help to identify the sources. It must be noted that although the PAO is not a dedicated neutrino experiment, it is capable of producing competitive results.

3.3.3.2 Magnetic Fields

In contrast to UHE photons and neutrino, cosmic rays suffer significant deflections caused by magnetic fields, especially at lower energies. A simulation of cosmic rays (10 protons each) of different energies flying through random magnetic fields (1nG) with a coherence length of 1 Mpc is shown in Fig. 3.9. For an energy of 1 EeV the complete source information is lost while for those at 100 EeV after 50 Mpc the source could still be identified, i.e. UHECR astronomy might be possible with these parameters.

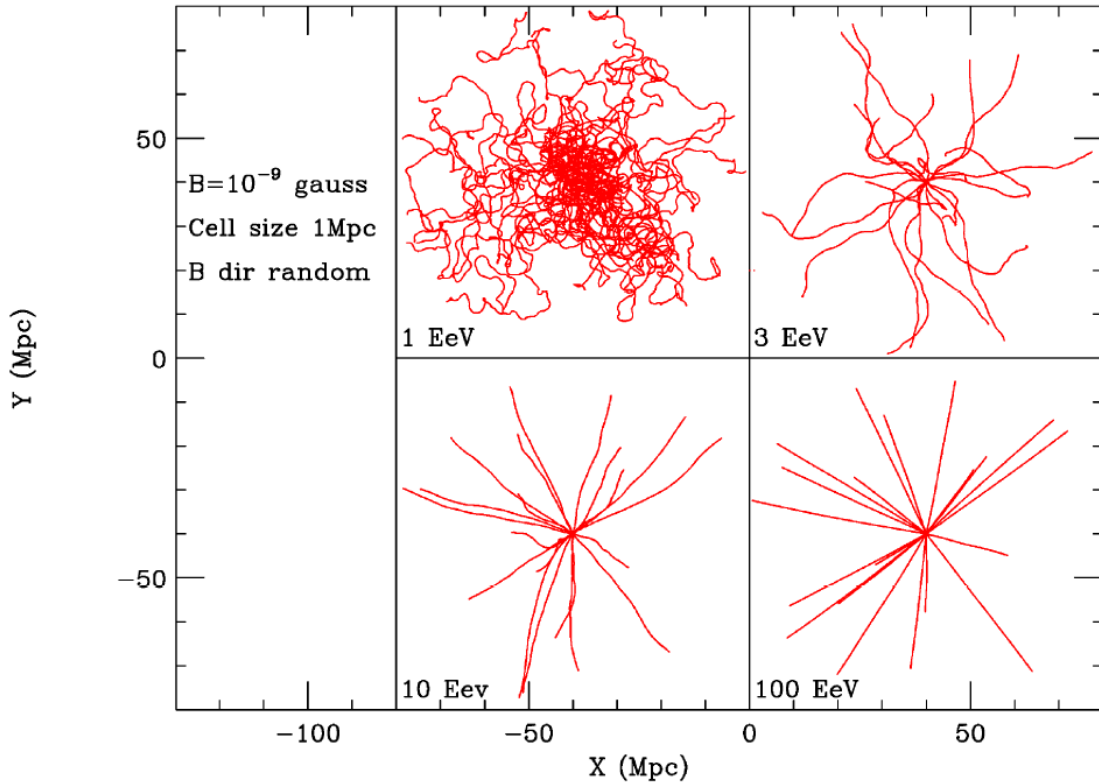


Figure 3.9: A simple simulation of UHECR (10 protons each) propagation with different energies, a random magnetic field of 1 nG with a coherence length of 1 Mpc. [41]

The above used values for the intergalactic magnetic field have been set recently as an upper limit by a study of the CMB using among other things Faraday rotation [42]. Furthermore, a new approach has been made in 2009 using UHECR data to measure the intergalactic magnetic field [43].

In general, it is assumed that the intergalactic magnetic field is randomly oriented, only in areas close to large matter distributions (e.g. galaxy clusters) filamentary structures with coherent fields are formed. Those accumulations can increase the strength of the field by a factor of roughly 100. In general, this assumption is called the random walk [44] model for UHECRs propagation since their deflection is leading to a Gaussian distribution for the arrival directions around a source.

Spiral galaxies like our own are known to have a large scale coherent regular field. This leads to more complex models of the field distributions. Two basic shapes are proposed:

- bisymmetric (*BSS*);
- axisymmetric (*ASS*).

In a *BSS* spiral galaxy the field vector has the same orientation for adjacent arms, whereupon in the *ASS* scenario they have the opposite sign.

An extra component can be added by making a distinction between the part above and below the galactic mid-plane. Are the two halves going in the same (opposite) directions then a S (A) is appended representing a symmetric (antisymmetric) behaviour.

The strength of a few μG is again estimated by Faraday rotation measurements of the plane of polarisation using pulsars in our own galaxy or extra-galactic radio sources. While for the highest energies the influence of the galaxy is considered to be negligible, UHECR with energies close to 10^{18} eV are significantly affected. These in turn might be used to estimate the strength and reconstruct the direction of the coherent magnetic field. This can be done for example by analysing the data in terms of energy ordered multiplets.

3.3.4 Low Energy Anisotropy Studies

In the case of lower energies ($E \gtrsim 10^{18}$ eV) also anisotropy studies are possible despite the effects of magnetic fields. The focus lies on large scale structures maybe caused by magnetic fields or large objects in our vicinity. Therefore, one concentrates here on multipole studies at larger angular scales, a standard method in this field of research, and objects within our galaxy.

In this sense, the PAO has formulated a prescription regarding an excess in the direction of our galactic center (GC) [45].

The data of a fixed period was scanned to find the parameter set which yields the largest deviation from the isotropic expectation. Then, one has two options:

- running prescription and
- fixed exit points.

While the first type of prescription ends whenever the requested confidence level is passed (cf. the AGN correlation study by the PAO), the second type has fixed ends. In advance, the points of exits are defined mostly depending on the number of events and if one of these numbers is reached, the data set is analysed using the pre-defined parameters. Here, the second approach is chosen.

The search for an excess in the direction of the GC can be motivated by two arguments. The first one is that the GC might be a lens for CR focusing them from the center towards the spiral arms of the galaxy. The second one refers to a conversion taking place within the GC. Protons are caught inside its magnetic field, combine with electrons or anti-neutrinos producing neutrons and neutrinos or positrons. This neutron can then escape the magnetic field and has enough energy to travel far enough in the direction of the Earth before it decays into a proton so that the deflection caused by the intergalactic magnetic field is smaller. Until now, the prescription has not yet reached the needed number of events to evaluate the test.

In principle, multipole studies could also identify the GC as a source whereupon dedicated searches need less statistics to be conclusive especially when it comes to large multipole coefficients l , i.e. small angular scales. Considering multipoles of lower

order (dipole, quadrupole), large scale studies are possible to identify an anisotropy at an early stage of an experiment. It has the advantage to be model independent, i.e. with this method extended regions in the sky can be found without previous assumptions about sources and their distributions. For example, UHECRs emitted by a single source nearby could still carry enough information about their true origin. Hence, this source could produce a dipole in arrival directions of detected cosmic rays with an amplitude of a few %. A similar signal can be caused by many weak sources and their UHECRs focused into one region of the sky by the coherent part of the galactic magnetic field.

In addition, the Compton-Getting effect [46],[47] can cause a dipolar structure in the arrival directions. If UHECRs are isotropic in our galaxy, then an earthbound experiment would see an increase in the rate of cosmic rays in the direction of motion. The same consideration applies also for the motion of the earth through the CMB rest frame resulting in an extra galactic Compton-Getting effect.

In the case of small amplitudes it is necessary to have a complete understanding of local effects. The geomagnetic field for example deflects the charged particles in an air shower depending on their relative orientation. This influences the energy reconstruction for surface detectors that sample the signal at ground level. Consequently, UHECRs of the same energy are assigned different energies by the air shower reconstruction depending on their arrival direction relative to Earth's B-field. This leads to an anisotropy in local coordinates when making a sharp energy cut. In turn a dipole is produced with its maximum at the south pole in equatorial coordinates with an amplitude of roughly 1%. To avoid this bias, a 1-dimensional harmonic analysis can be performed in right ascension α because then each bin in α is affected in the same way. A corresponding analysis by the PAO has recently been published [48]. In this context, a procedure was applied which allowed to use air showers reconstructed with an energy below the lower threshold of $3 \cdot 10^{18}$ eV for the full acceptance of the experiment. No significant amplitude has been found so far, only limits have been set for energies larger than $0.25 \cdot 10^{18}$ eV (see Fig. 3.10). In contrast to these results, AGASA reported an amplitude of 4 % in one energy bin in [49]. One reason for this discrepancy might be the different hemispheres the experiments are located.

3.3.5 Autocorrelation Studies

By adapting the concept of model-independent searches for anisotropy at the highest energies, autocorrelation studies offer the possibility to recognize anisotropies already in the case of low statistics. Thus, the arrival directions of UHECRs can be assessed at the highest energies. This has the advantage of being completely free of any assumptions or preconceived opinions about sources or at least their distributions. In this way, new hints for the true origins can be gained.

A first approach was made by AGASA searching in their data set for a clustering on small scales [50]. With a minimal energy of $40 \cdot 10^{18}$ eV an over-abundance of doublets and triplets was discovered close to the super galactic plane.

A similar study was carried out by the PAO. After the positive anisotropic signal from the correlation analysis was obtained, further efforts were made to establish

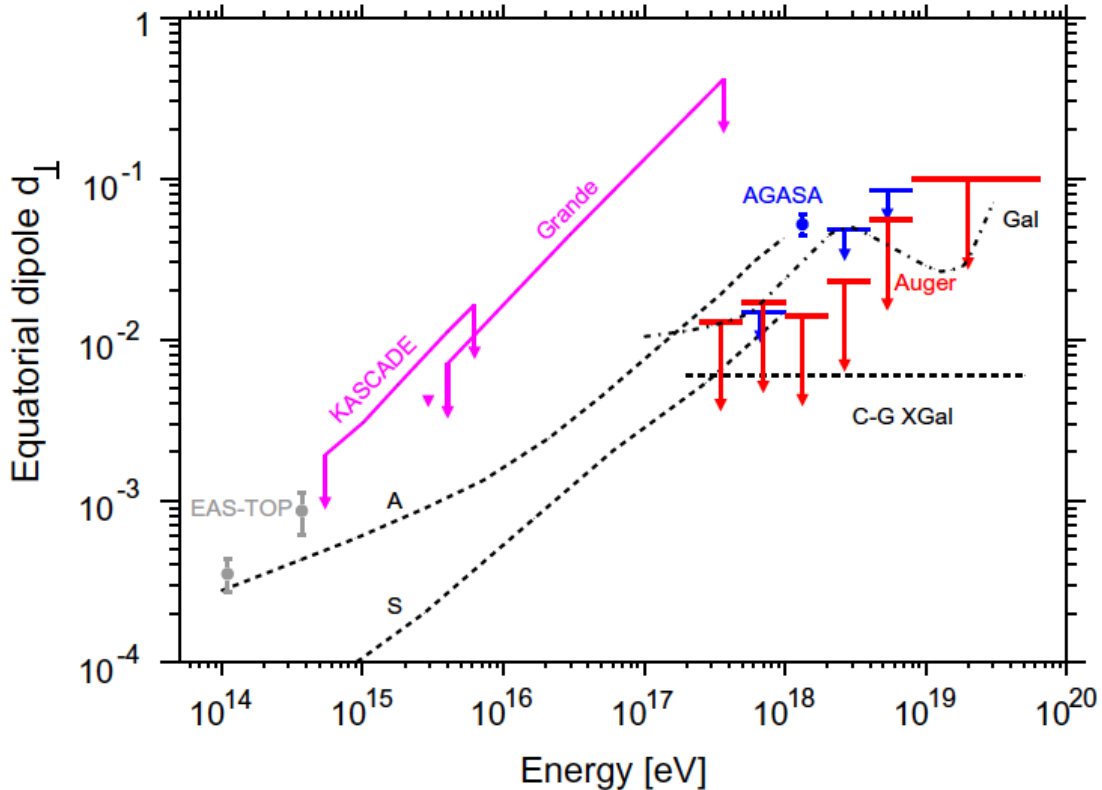


Figure 3.10: Upper limits on the anisotropy amplitude of first harmonic as a function of energy in right ascension. Results from EAS-TOP, AGASA, KASCADE and KASCADE Grande experiments are displayed too. Also shown are the predictions up to 1 EeV from two different galactic magnetic field models with different symmetries (A and S), the predictions for a purely galactic origin of UHECRs up to a few tens of 10^{19} eV (Gal), and the expectations from the Compton-Getting effect for an extragalactic component isotropic in the CMB rest frame (C-GXgal). [48]

a consistent result using autocorrelation methods. An a posteriori analysis in 2009 [51] showed that the optimal lower energy cut of $52 \cdot 10^{18}$ eV for this is close to the one obtained by the scan in the run-up to the prescription in 2006. The minimal not penalised chance probability P_{chance} is 0.26 %.

As a reminder, on this topic the main focus of this work is put. Therefore a more detailed discussion about assets and drawbacks as well as developed methods in this context is carried out in chapter 5.

3.4 Extensive Air Showers

All the measurements of UHECRs and corresponding theories described for energies $E > 10^{15}$ eV rely on the detection of extensive air showers. Air showers consist of a multitude of secondary particles and are initiated by primary particles interacting with nuclei in the atmosphere. These secondary particles in turn interact with other nuclei producing even more particles generating a cascade of reactions. Different detector types can then measure the longitudinal and lateral shower profiles (see

chapter 3.4.2.1 and 3.4.2.2).

Especially for detectors sampling and reconstructing the signal at ground level it is essential to know which part of the evolving shower is detected. For example, if two primary particles with the same energy but different arrival directions penetrate the atmosphere and have their first interaction at different heights, the number of particles and hence, the strength of the measured signal would be different. In this context, Monte Carlo shower simulations are an important tool to interpret the air shower measurements. Due to the underlying statistical processes, even showers with the same initial properties will have a different shower development. Therefore, an analysis on an event-by-event basis is impossible. A sufficient amount of statistics must be gathered to draw a conclusion.

To this end, it is important to understand the underlying processes in the shower and in addition, to know the values of involved variables, e.g. of the coupling constant of the strong force, α_S , and of the proton-Nucleus cross-section, σ_{pN} . Although, current accelerators are running at energies of the order of 10^{12} eV and experiments such as CM or ATLAS are measuring with a precision never reached before, a transfer to the field of physics of cosmic rays poses difficulties. For instance, at the LAC mainly proton-proton collisions are studied by means of the corresponding reaction products flying perpendicular to the beam pipe. In contrast to that, air shower experiments are interested in the collisions of protons and heavier nuclei (nitrogen and oxygen) as well as in the reaction products flying nearly parallel to the beam pipe. Therefore, a simple model predicting accurate properties of a shower is advantageous and will be presented in the following subsections. Furthermore, we discuss two detection techniques of extensive air showers.

3.4.1 Shower Development

In general, four components of a shower are distinguished:

1. electromagnetic component consisting of electrons, positrons and photons;
2. hadronic component consisting of mesons and baryons;
3. muonic component;
4. invisible component consisting of neutrinos.

While electromagnetic (EM) primary particles will generate pure EM showers, a hadronic induced EAS, however, will produce all components.

We start with the simple case of an electromagnetic shower, the original Hitler model, and continue with an adaption towards hadronic showers following the argumentation presented in [52].

3.4.1.1 Electromagnetic Shower

In the 50s of the last century Heitler developed a method to describe analytically the development of an electromagnetic air shower in the atmosphere. Initiated by a

photon at the first interaction point, the cascade starts with the production of an e^+e^- -pair transferring the complete energy equally to both particles. After a fixed distance, also referred as slant depth $d = \lambda_r \ln 2$, they in turn emit bremsstrahlung whereupon again the energy is divided equally between photon and e^\pm (cf. left panel Fig. 3.11). d depends on the radiation length λ_r of the particle in air. The factor $\ln 2$ ensures that d corresponds to the typical distance after which electrons lose half of their energy by radiation. The distance is measured in units of $g \text{ cm}^{-2}$ indicating how much material per area has been traversed by the shower. Hence, $d = 0 \text{ g cm}^{-2}$ corresponds to the top of the atmosphere.

In the Heitler model of the cascade process, e^\pm and photons emit bremsstrahlung and produce e^\pm -pairs, respectively, after each interaction step of distance d . This increases the number of particles N by a factor of 2 after each step. Consequently, after n steps the traversed distance is $x = n \cdot d$ and $N = 2^n$. These reactions continue until the energy per particle is below the threshold energy E_{thr} for the processes mentioned above. Afterwards, losses due to collisions exceed those caused by radiation. At that energy, the shower contains the maximal number of particles $N_{max} = 2^{n_{thr}}$ with energy $E_{thr} = E/N_{max}$ each. This corresponds to a slant depth of

$$X_{max} = n_{thr} \lambda_r \ln 2 + C = \lambda_r \ln(E/E_{thr}) + C, \quad (3.4)$$

with C the height of the first interaction. With these results the simple Heitler model makes two correct predictions regarding extensive air showers initiated by a photon: On the one hand, the number of particles increases linearly and on the other hand the depth of the shower maximum logarithmically with energy.

Despite the good performance, there are several points where the predictions of the model deviate from measurements. For instance, the numbers of e^\pm is overestimated by roughly a factor of 10. One reason for this is the simplification in the Heitler model assuming a straight line of flight for the electrons, whereas they in fact spread out laterally and are attenuated while flying through the atmosphere.

3.4.1.2 Hadronic Showers

At the highest energies, however, cosmic rays are dominated by nuclei. Therefore, a similar model was developed to analytically describe extensive air showers initiated by hadrons. Here, we discuss the case of a proton with $E = 10^{18} \text{ eV}$. The first alteration affects the multiplicity after one interaction length (λ_I). In contrast to electromagnetic showers, after each step the number of particles is increased by a factor of approximately 10. Another important factor is related to the type of the secondary particles: In one interaction a number of N_{ch} charged as well as N_0 neutral pions is produced whereupon each type has the same probability. The resulting relation of charged to neutral is $N_{ch} = 2 \cdot N_0$.

A π_0 is very short lived even with velocities close to the speed of light in the frame of reference of the laboratory. It decays immediately exclusively in two photons, although a small branching ratio exist for other decay modes. The two photons start an electromagnetic air shower each following the description from above. Charged pions (π^+ and π^-), however, fly the average interaction length λ_I before they react again with nuclei of nitrogen or oxygen producing more pions in the same ratio (cf.

right panel of Fig. 3.11). Consequently, the number of pions increases with each step, $N_\pi = (N_{ch})^n$ with an energy per particle of $(2/3)^n E$.

After several similar interaction steps n_{thr}^π , the energy per charged pion drops below a threshold E_{thr}^π where the probability to decay is becoming dominant compared to the probability to fly the whole interaction length λ_I . In this model, all pions decay into muons ($N_\pi = N_\mu$) and neutrinos at this point. The formula for this energy is:

$$E_{thr}^\pi = E \cdot \left(\frac{3}{2}N_{ch}\right)^{-n_{thr}^\pi}. \quad (3.5)$$

Solving the equation to obtain the number of interactions gives

$$n_{thr}^\pi = \frac{\ln(E/E_{thr}^\pi)}{\ln(\frac{3}{2}N_{ch})}. \quad (3.6)$$

To assess the energy of the primary particle also the electromagnetic part has to be included, $E = E_{thr} \cdot N_{max} + E_{thr}^\pi \cdot N_\mu$. Both threshold energies are estimable in principle offering the possibility to determine the total energy of the primary particle by measuring the amount of electrons and muons at the maximal size of the shower. Regarding the depth of the shower maximum, only the electromagnetic component is considered. It forms the major component of a hadronic air shower since $\lambda_r < \lambda_I$. Due to simplicity only the first π^0 decay contributes. At the beginning of the electromagnetic cascade the same amount of photons and charged pions is present ($\pi^0 \rightarrow \gamma\gamma$). These photons produce the fourfold amount of electrons within two radiation lengths while the number of pions is constant. Considering the same ratio for further interaction and radiation lengths respectively, electrons are the dominant component of the hadronic EAS. Therefore, we can use the formula from the pure electromagnetic case accounting for the point of the first interaction of the proton by adding an offset C . Here it is assumed that all photons carry an energy of $E/(3N_{ch,thr})$ in total,

$$X_{max}^\pi = \lambda_r \ln[E/(3N_{ch,thr})] + C, \quad (3.7)$$

giving the same logarithmic dependency of the primary energy.

As before, this model is incomplete containing similar simplifications as in the electromagnetic model. For example, many more types of particles are created such as charged and neutral kaons contributing to different parts of the shower. In addition, the electromagnetic cascade is fed by further decays of neutral pions produced after each interaction length.

The model can be extended to heavier nuclei which is not discussed in this work.

3.4.2 Detection Techniques of EAS

After presenting two simplified analytical models of the development of air showers, we now discuss the techniques how to measure the shower properties to obtain the results presented earlier in this chapter. There are two ways to determine the nature of the primary particle by measuring:

1. the longitudinal shower profile and/or

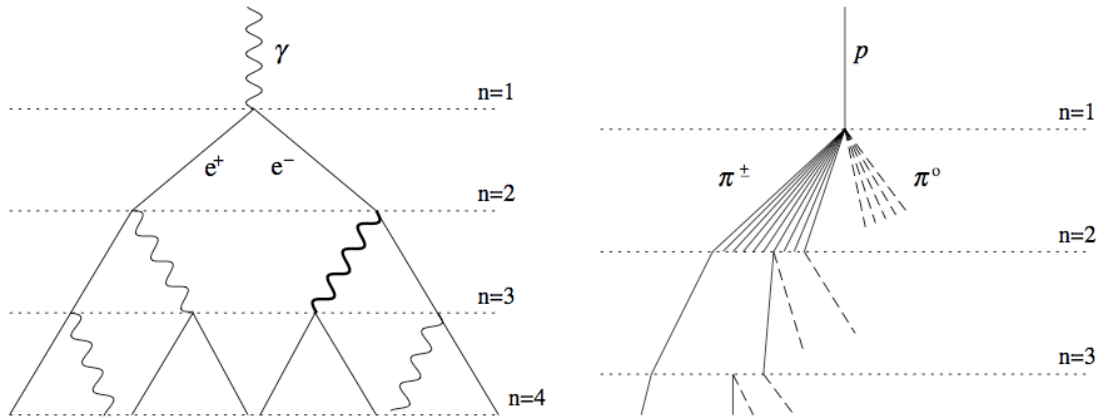


Figure 3.11: Schematic Extensive Air Shower development. *Left Panel:* An electromagnetic cascade is shown. After each interaction point n the number of particle doubles. *Right Panel:* A simplified hadronic cascade; not all particles are tracked through all interaction steps. Both taken from [52].

2. the lateral shower profile.

A sketch of the detection of air showers is shown in Fig. 3.12.

3.4.2.1 Measuring the Longitudinal Shower Profile

During an EAS, the particles travel through the atmosphere and the electrons cause two kinds of radiation. The first one is Cerenkov light emitted by the particles when their velocity exceeds the speed of light in the respective medium. It is emitted in the direction of movement. Hence, due to the small flux of UHECR and the opening angle which is very small for velocities close to c , an array of Cerenkov telescopes would be too inefficient to observe a sufficient amount of events in a reasonable time. In contrast to that stands the second kind, the fluorescence light, which is isotropically radiated by excited nitrogen. It offers the possibility to observe and study the longitudinal shower development in more detail covering a huge atmospheric volume to compensate the small flux. Since, here the atmosphere works as a calorimeter an accurate monitoring is required. This includes among other things, temperature and density of the air. The fluorescence yield of nitrogen depending on these factors must be measured. Without these reference parameters a shower reconstruction for instance leads to larger uncertainties.

A huge mirror collects the dim fluorescence light on a camera consisting of many Photo Multiplier Tubes (PMTs). The track projected on these pixels (=PMTs, see for an example Fig. 3.13) is used to reconstruct in the first place direction and energy.

For this purpose, on the one hand the timing information is needed to reconstruct the direction and on the other hand the signal strength in each PMT for the energy. The strength is proportional to the amount of fluorescence light produced at a certain slant depth X in the atmosphere which is directly connected to the number of charged particles N_e within the shower at that position. The observed longitudinal

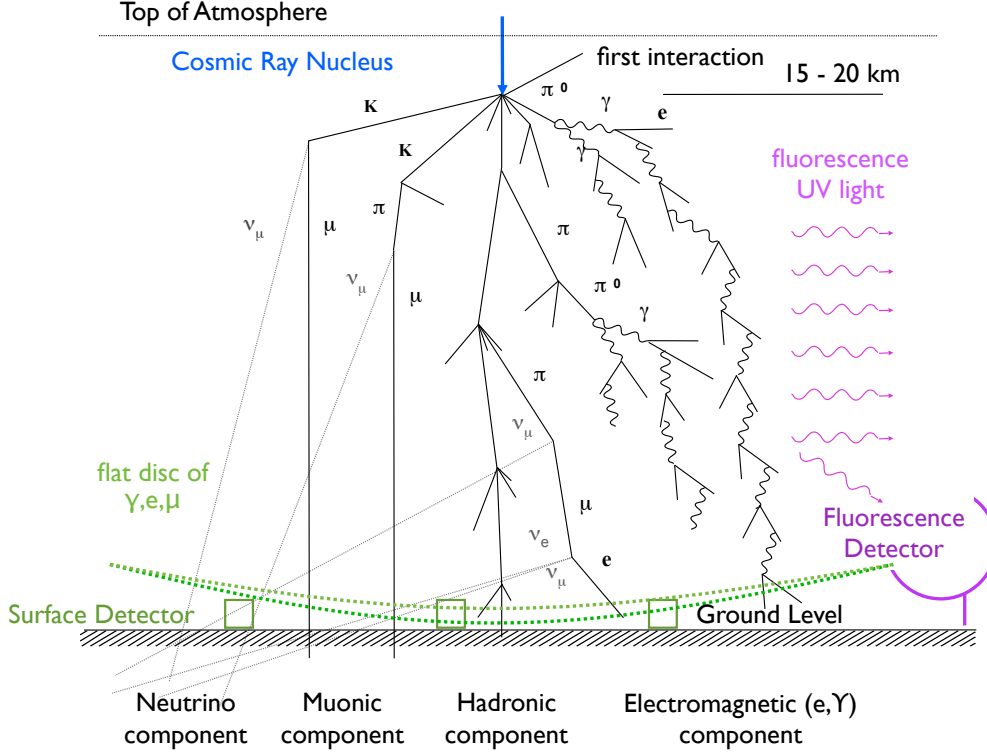


Figure 3.12: Scheme of detection principles of Extensive Air Showers. A nucleus penetrates the atmosphere and interacts with the atoms within. A cascade of secondary particles develops consisting amongst other of muons and electrons. While traversing the atmosphere, nitrogen atoms are excited which in turn lose this energy by emitting isotropically fluorescence light. This is detected by the corresponding detector measuring the longitudinal shower profile. Particles reaching the ground are used to sample the lateral shower profile. (adapted from H. Dembinski)

profile can be parametrised with the Gaisser-Hillas-function describing the development of electron numbers from the first interaction until the shower fades out.

$$N_e(X) = N_{max} \left(\frac{X - X_0}{X_{max} - X_0} \right)^{\frac{X_{max} - X_0}{\lambda_I}} \cdot e^{-\frac{X_{max} - X}{\lambda_I}}, \quad (3.8)$$

with N_{max} the maximal number of shower particles reached at a slant depth of X_{max} . X_0 is the slant depth of the first interaction and λ_I the assumed interaction length. An example is given for one event of the PAO in Fig. 3.14 showing a good agreement between model and measurement. A fit of the data with this formula yields the energy within the shower as well as the slant depth where the shower reaches its maximum, X_{max} . The first can be deduced by integrating over the complete range of X ; it must be scaled by multiplicative factor determined by the threshold energy of electrons in air divided by the radiation length in that very medium [53]. The determination of X_{max} depends on the knowledge about the point of the first

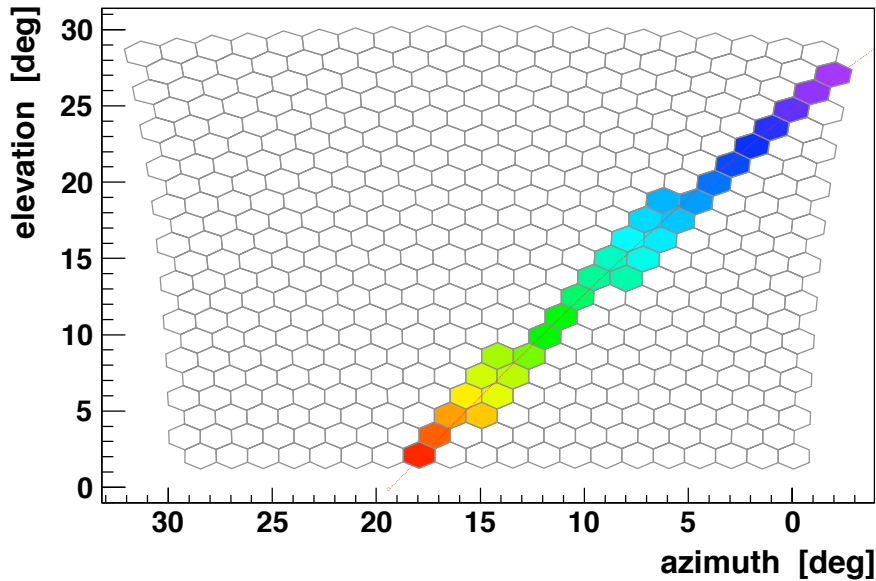


Figure 3.13: An example track through the pixels of fluorescence telescope of the PAO is shown. The colour code represents the timing (purple = early, red = late). The thin red line represents the plane which contains the air shower and a reference point of the telescope.

interaction X_0 which varies even for primary particles with the same energy coming from the same direction. The error caused by this fluctuation is of the order of 10 %. From the study of the corresponding distribution of X_{max} , the composition can be estimated, since $X_{max}(Fe)$ is smaller than $X_{max}(p)$ for the same energy.

Next to the listed advantages of EAS measurements, there exist also some drawbacks for fluorescence telescopes. For example, this detection technique relies strongly on a very small amount of background light. Hence, only moonless nights and locations with a minimal contribution of diffuse light are allowed. Otherwise, signals caused by UHECRs would be undetectable.

In addition, local conditions like clouds and air pollution can also restrain the measurement by absorbing parts of the fluorescence light of the shower. Including all these effects, the duty cycle is of the order of 10 %. However, experiments like HiRes exclusively used this technique and delivered competitive results with the PAO, for instance.

3.4.2.2 Measuring the Lateral Shower Profile

A complementary approach is the sampling of the shower particles at ground level. To this end, a huge area has to be instrumented by many spatially separated detectors. The distance between the individual stations determines the detection efficiency and its dependence of the energy of the UHECR. For the highest energies, a spacing of more than 1 km is appropriate. An EAS reaching the Earth is again reconstructed using time and signal strength information in each detector.

The first approach assumes a flat shower front of secondary particles to have a first

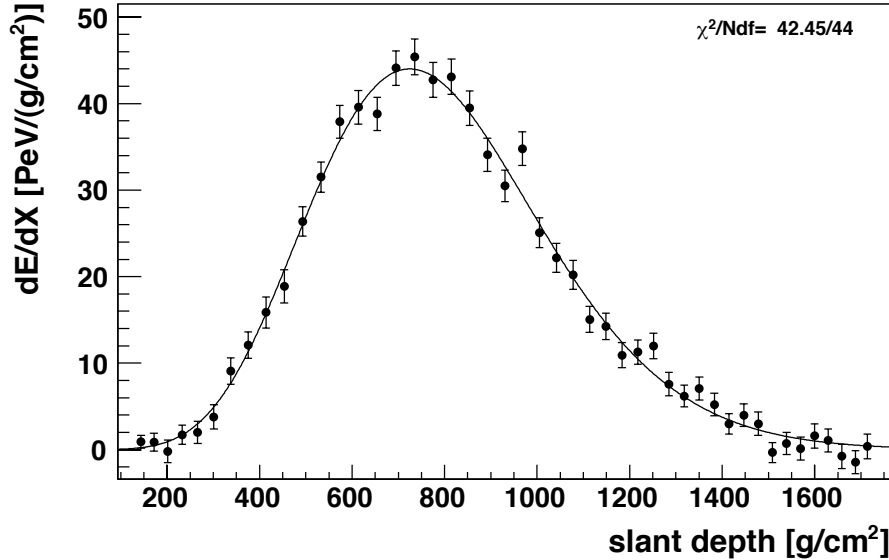


Figure 3.14: A Gaisser Hillas Fit to an event measured with the fluorescence telescope of the PAO (taken from [54]).

estimate of the shower axis and hence, of the shower core. The signals are projected to a plane perpendicular to the the shower axis. After determining the shower core in this plane, a function formulated by Nishimura, Kamata and Greisen (NKG) describing the lateral spread density, $\rho(r)$, of an EAS allows to estimate the total number of particles hitting the ground. ρ depends on the distance r of the signals in the shower plane to the shower core. An azimuthal symmetry around the shower core is assumed.

$$\rho(r) = \frac{C_1(s)N_e}{2\pi r^2} \left(\frac{r}{r_1}\right)^{s-2} \left(1 + \left(\frac{r}{r_1}\right)\right)^{s-9/2} \left(1 + C_2 \left(\frac{r}{r_1}\right)^\delta\right) \quad (3.9)$$

Here, C_i are normalisation constants and N_e the expected number of electrons on the ground of an air shower. r_1 is the Moliere radius defined in a way that within a radius of $2r_1$ 95 % of the energy is deposited. Furthermore, the exponents s and δ are estimated from simulations and have to be adjusted for each experiment.

An alternative determination of the energy is based on the particle density measured at a fixed distance from the shower core. A reasonable choice verified by simulations has the advantage of being independent of the nature of the primary particle. For example, the AGASA experiment fixed this distance to 600 m. Again, this number depends on the experiment and its properties, e.g. distance of stations relative to each other, and altitude. An example of the lateral distribution of signal measurements of a real event detected at the PAO is shown in Fig. 3.15. An important advantage of this technique is the 100 % duty cycle, it is essentially independent of weather conditions. However, the strong dependence on simulations and thus on interaction models increases the uncertainty on the energy estimator.

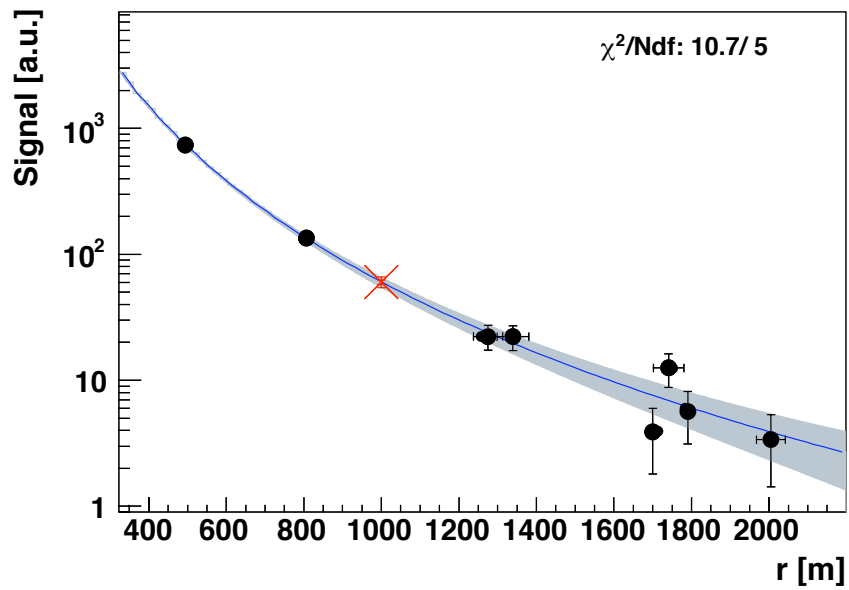


Figure 3.15: A NKG Fit to an event measured with the surface detector of the PAO. The circles are the signals of a station with a certain distance to the shower core. In grey shade the NKG fit including error band is shown. The red cross marks the point in the lateral distribution used to estimate the energy here at 1000 m away from the shower core. The signal on the y-axis is given in arbitrary units.

4. The Pierre Auger Observatory

Until the beginning of this century solely experiments focussing on one detection principle described in the previous chapter were and have been active :

- fluorescence telescopes (e.g. HiRes, sensitive to the longitudinal profile) and
- surface detectors (e.g. AGASA, sensitive to the lateral profile).

Therefore, a new large observatory was proposed combining both techniques. It should exceed all predecessor experiments in size allowing to study UHECRs with a level of precision never reached before at energies above $3 \cdot 10^{18}$ eV.

In 2004, the Pierre Auger Observatory (PAO) close to Malargüe, Argentina, fulfilling this requirement started data acquisition on a regular basis while still being under construction. It is located in the Argentinian pampa at an altitude of 1400 m where the landscape is rather flat so that the reconstruction of events is less influenced by geographical circumstances.

Four years later it was completed covering in total an area of 3000 km² with 1600 water Cerenkov stations forming the surface detector (SD). This area is overlooked by the fluorescence detector (FD) consisting of four buildings called eyes. Due to the combination of FD and SD, the PAO is defined as a hybrid detector profiting from the possibility to cross-calibrate between the two detector types. For instance, the uncertainty on the energy reconstruction of SD events can be reduced by calibrating it with FD which has a better energy resolution based on its direct calorimetric measurement of the longitudinal shower profile. Otherwise, the results obtained from the analysis of SD data would be less accurate relying on air shower simulations. To obtain the corresponding calibration function, events are used which have been detected and reconstructed separately by FD as well as SD called golden hybrids. This example shows, how the downside of one detector type can be compensated by the other.

The general idea behind the PAO was and is to have two sites, one on the northern and one the southern hemisphere to cover the whole sky with one instrument. Until now, only the above described southern part is completed. This is one reason why some discrepancies between experiments could not be resolved, e.g. the differing composition measurements of HiRes and the PAO. An observatory with the same properties on both hemispheres of the Earth could easily answer the question whether the discrepancies originate from the choice of the site or depend on the reconstruction algorithms and detector specifications.

In the following both detector types are described; the focus lies on SD, since the data analysed in this work is based on pure SD measurements. Subsequently, enhancements to the existing PAO are discussed. Particularly interesting is the High Elevation Auger Telescope consisting of three inclineable shelters. We discuss in detail how the optical system is monitored to ensure a stable performance.

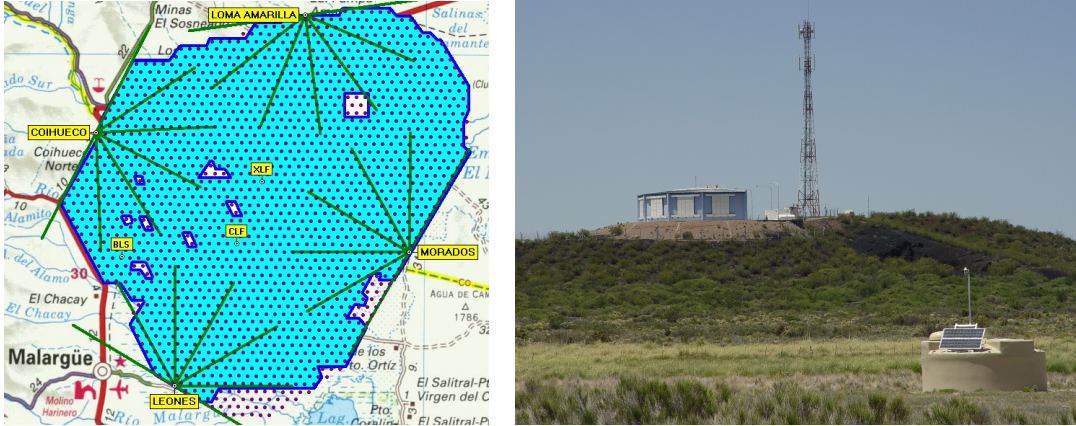


Figure 4.1: *Left Panel:* A map of the Pierre Auger Observatory covering with the surface detector (SD, red dots) an area 3000 km^2 (blue shaded). The fluorescence detector (FD) consisting of four sites, Los Leones, Los Morados, Loma Amarilla and Cuihueco overlooks the array. *Right Panel:* A picture of an FD building with an SD station in front.

4.1 Fluorescence Detector

The fluorescence detector (see [54]) consists of four buildings placed in all cardinal points covering the array of the surface detector (cf. left panel of Fig. 4.1). It uses the calorimetric information provided by the atmosphere and the timing of the signal to reconstruct the properties of an EAS as discussed in chapter 3.4.2.1.

Each of these buildings houses six telescopes shown in Fig. 4.2 with a field of view of 28.1° above the horizon and 30° in azimuth covering in total 180° in azimuth. A telescope consists of a $3.5 \text{ m} \times 3.5 \text{ m}$ spherical mirror focusing the gathered light onto a camera of 440 pixels arranged in a 20×22 matrix. Each pixel is a PMT with a hexagonal front and a field of view of 1.5° .

In addition, a filter is included which is transparent mainly in the ultraviolet band suppressing the amount of background light at differing wavelengths. Note that the emitted fluorescence light contributes primarily to the UV-range.

To estimate the energy of the primary particle, it is necessary to know the fluorescence yield in air, i.e. the number of emitted photons per deposited energy. The amount of light reaching the PMTs depends on atmospheric conditions. Hence, atmospheric distortions has to be taken into account leading to the requirement to monitor constantly the atmosphere. Due to the importance of the fluorescence yield for the energy determination, dedicated experiments exist such as the AIRFLY experiment [55].

Atmospheric effects like thunderstorms and the restriction of data taking to moonless nights reduce the duty cycle of the fluorescence detector to 13 % in the end.

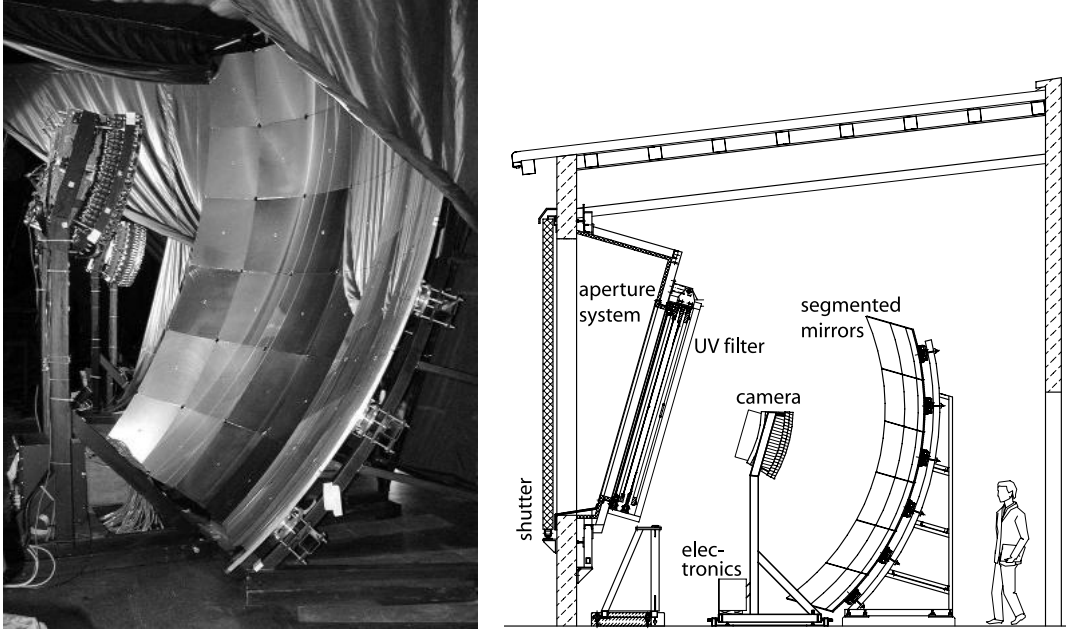


Figure 4.2: The fluorescence detector of the Pierre Auger Observatory. *Left Panel:* A photo of one camera including the mirror within a FD building. *Right Panel:* A schematic sketch of a bay within an FD building including the shutter, the UV-filter, mirror, camera and electronics in the DAQ box.

4.1.1 Reconstruction Principle and Resolution

The first two trigger stages within an FD telescope are implemented in the hardware. While the first level trigger (FLT) considers only single pixels and checks whether they exceed a certain signal threshold, the second level trigger (SLT) processes the structure of the track through the camera. Within a detected structure the algorithm searches for five basic shapes (cf. Fig. 4.3) including their rotations. However, a minimal number of five pixels is required to pass the SLT. In addition, if a track is recorded and only one pixel is missing which would complete one of the five shapes, then the SLT also passes. Subsequently, the found tracks are analysed by a software to reject background events, e.g. those caused by lightning. The efficiency here is 94 %, while only 0.7 % of true events are thrown away.

For the reconstruction of the EAS additional variables need to be defined. The track measured by one or more telescopes in one eye (see left panel of Fig. 4.5) determines the shower-detector plane (SDP). The triggered pixels (purple stands for early, red for late) correspond to a fixed azimuth angle and elevation. By following the evolution of the shower through the sky and thus, through the pixels, these connections form, together with a point representing the FD building, the SDP (cf. Fig. 4.4) with an inclination of ϑ_{SDP} . In addition, for each pixel the viewing angle χ_i is calculated by means of the pointing information of the triggered pixel. It is the solid angle between the ground of the array and the position (azimuth, elevation) of

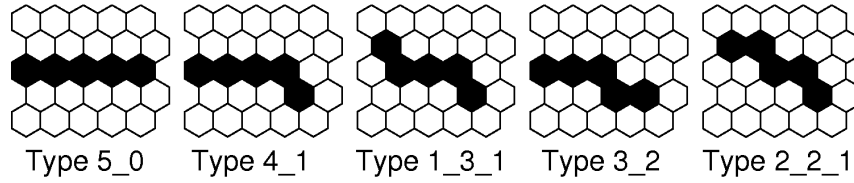


Figure 4.3: These are the fundamental types of patterns regarded as straight track segments for the second level trigger of FD.

the pixel within the SDP.

By using the timing information t_i of a pixel as well as the corresponding χ_i under the assumption that the shower propagates with the speed of light, the shower axis and the distance to the telescope can be fitted by means of a χ^2 minimization. This is called a mono reconstruction considering only one eye.

Further information can be gained by including the timing and viewing angle χ_i of a triggered SD station. The station with the largest signal commonly known as the hottest station (black square in Fig. 4.5) pins down the impact point of the shower on the ground. As a result, the errors on shower core position and shower axis reconstruction are reduced significantly, especially in cases of tracks with only few triggered pixels. Therefore, these events, commonly known as hybrid events, are preferred over mono events.

In general, after a physical event was recognized by the fast FD reconstruction software directly after detection, the SD array in front of the eye is read out increasing the probability to obtain a hybrid event. Note that a single triggered SD tank remains normally unrecorded (see below).

Events enabling a hybrid as well as a pure surface detector reconstruction are called golden hybrids. These are the ones fitting for the cross-calibration procedure which is discussed in the next section.

After the geometrical shower properties are set for an EAS, the signal strength in the individual PMT is considered. As discussed in chapter 3.4.2.1, applying a Gaisser-Hillas-Fit to these signals gives the value of X_{max} , the slant depth where the number of particles of the recorded EAS is maximal. In addition, during the reconstruction algorithm further radiation constituents are calculated, e.g. the contribution of direct and indirect (scattered) Cerenkov radiation. This is particularly important for those showers traversing the atmosphere towards the telescopes. Cerenkov radiation is emitted in forward direction and contributes therefore to the signals in the PMTs. Regarding the reconstruction accuracy of energy, arrival direction and X_{max} , all the estimations based on hybrid events rely on simulations [56],[57]. In the following the event-by-event resolutions are presented as obtained from simulations in an energy range of $10^{18} - 10^{19}$ eV:

$$\begin{aligned} \text{Energy Resolution:} &= 8\% \\ \text{Angular Resolution:} &< 1^\circ \\ X_{max} \text{ Resolution:} &= 20 \text{ g cm}^{-2}. \end{aligned}$$

In addition, the energy scale has a systematic uncertainty of 22 % mainly caused by the uncertainty on the fluorescence yield.

Studies with stereo events (seen by two or more eyes) and no SD station show comparable results.

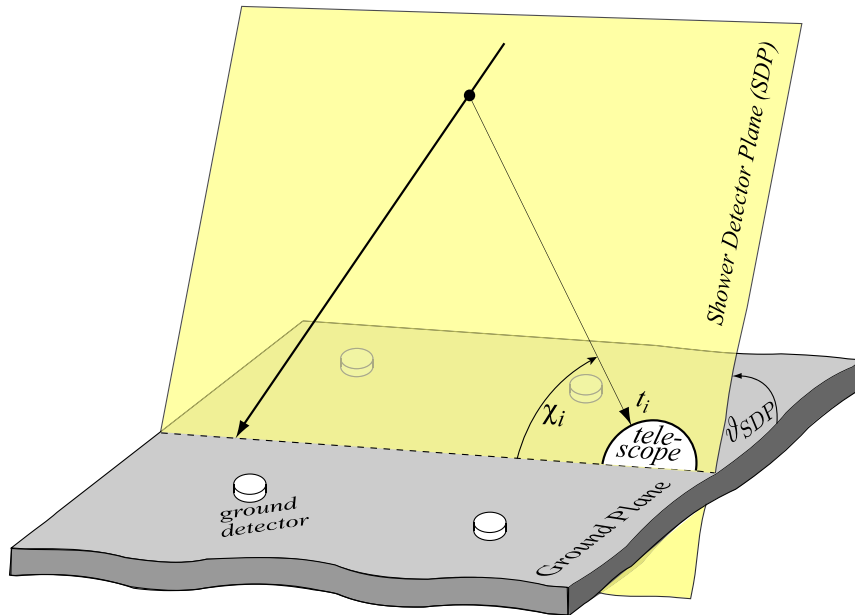


Figure 4.4: Geometrical shower properties of an event measured with the fluorescence detector. The shower axis and the telescope lie within the shower detector plane (SDP) which has an inclination of ϑ_{SDP} (adapted from [58]).

4.2 Surface Detector

The 1600 surface detector stations are arranged in a hexagonal grid with a spacing of 1.5 km. This array samples the lateral distribution of the secondary shower particles as described in chapter 3.4.2.2. Each station (cf. Fig. 4.6) consists of a cylindrical tank with a height of 1.2 m and a ground area of 10 m² containing 12 m³ of purified water within a highly reflective liner. While charged particles traverse the water in the station, they emit Cerenkov radiation since their velocity is larger than the speed of light in that medium. The light is reflected from the bottom as well as the top and the walls reaching in the end the three PMTs embedded in the upper shell looking into the water. The PMTs and corresponding electronics are powered by a battery which is charged by a solar panel. Thus, each station is self-sustaining, it needs only maintenance in case of failures. For position and timing determination a GPS antenna is mounted on top together with a radio antenna used for communication with the central data acquisition centre.

In the following we focus only on not very inclined events with a zenith angle smaller than 60° which follow the standard reconstruction procedure. Events with a larger angle need a different treatment due to their longer way through the atmosphere: the electromagnetic component of the shower dies out and the muonic one dominates the lateral distribution leading to a differing reconstruction procedure.

4.2.1 Trigger Structure

The trigger structure of the SD array consists of two phases [59]:

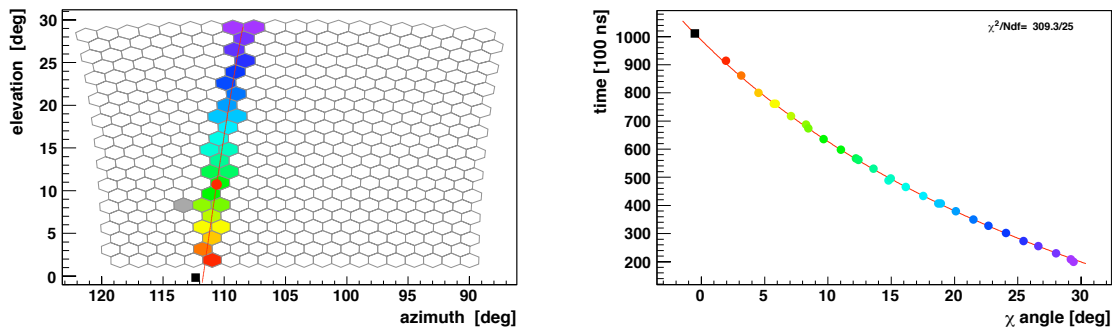


Figure 4.5: An example hybrid event with the energy $E = 2.43 \cdot 10^{19}$ eV seen by telescope 4 in Los Morados. The colour code in both pictures contains the time information, i.e. purple means early and red late. The black square represents the hottest SD station. *Left Panel:* Track of the event along the camera whereupon the red line is the shower detector plane and each hexagon represents one pixel of the camera. *Right Panel:* Viewing angle and timing of the triggered pixels are displayed including the best fit for all parameters describing the shower.

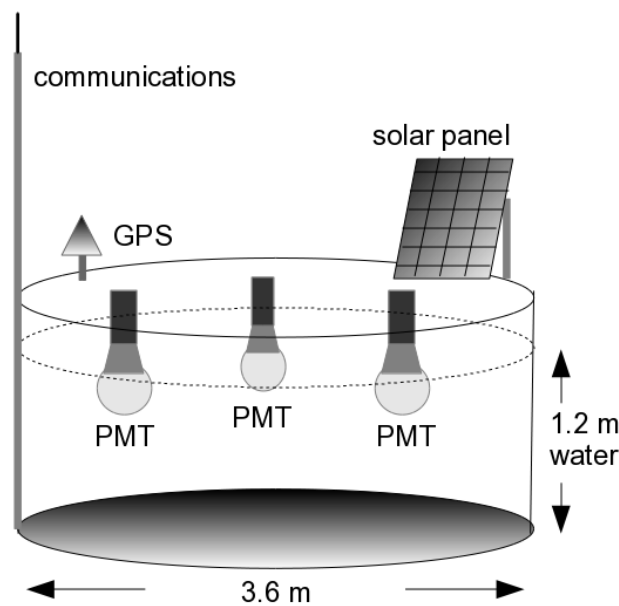


Figure 4.6: A sketch of a surface detector station of the PAO including the water tank with 3 PMTs embedded in the upper shell. On top are the GPS and communication antennas as well as the solar panel to supply power to the battery powering the electronics. (provided by A. Nelles)

1. station triggers (T1, T2) and
2. central data acquisition triggers (T3, T4, T5).

The T1 trigger can be fulfilled in two different ways. On the one hand, the pulse height, can cross a certain threshold (TH-T1) in all three PMTs in a station simultaneously to proceed further. On the other hand, a time-over-threshold (ToT-T1) is also allowed. Here the signal has to be constantly above a fixed value for more than 300 ns. This threshold value is smaller than the one used for the TH-T1. In case of PMT failures, the thresholds are adjusted, i.e. the fewer the number of active PMTs the higher the values have to be to reduce the background caused by electronic noise. The unit of the signal is chosen to be VEM which stands for vertical equivalent muon. From the uniform flux of atmospheric muons an average pulse height can be estimated which is set to 1 VEM above the electronic noise. All measured signals are given with respect to this value.

While the TH-T1 has to pass an additional higher signal threshold in conjunction with the T2 trigger before it is send to the central data acquisition system (CDAS) in Malargüe, the ToT-T1 is directly passed to CDAS.

In the next step (T3), the single T2 triggers are gathered according to their timing information. Here, again two types of events are distinguished, those including a TH-T1 and pure ToT-T1 triggers. Regarding the first case, among all T2 triggered stations four must be arranged as can be seen in the right panel of Fig. 4.7. Around a station the others are allocated in crowns C_n with n indicating the corresponding order. Here, at least one station in the first, second and fourth crown has to have a T2 based on TH- and/or ToT-T1. This mode is called $2C_1 \& 3C_2 \& 4C_4$. Considering the pure ToT-T2 case an additional configuration is allowed, namely the one displayed in the left panel of Fig. 4.7. It is comparable to the $2C_1 \& 3C_2 \& 4C_4$ whereas the triggered station in the fourth crown is omitted. To indicate its exclusive application to the ToT-T2 it is referred to as $ToT2C_1 \& 3C_2$.

If the spatial requirements are fulfilled, the timing has to be reasonable therefore the included stations have to be closer than $(6+5C_n) \mu s$ in their time difference to the central station.

To separate real events from background a more strict timing condition is applied within the T4. For pure ToT triggers within a T3 event, a triangle of triggered stations has to be found which is compatible with a flat shower front moving with the speed of light. Afterwards, further stations are added to check whether their time information agrees with the first estimation of time predicted by the shower front. If not, they are removed from the event and flagged as accidental which is also true for tanks which are farther away than 3 km from another triggered station.

T3 events including TH-T1 trigger, however, only need to have four stations passing the T2 without any restrictions regarding the spatial distribution, i.e. no triangular structure is needed. The rest of the procedure is the same as before.

The last step, T5, is a fiducial trigger. To pass it, the hottest tank within an event has to be surrounded by six working tanks in the first crown C_1 , $6T5$. This ensures a reasonable reconstruction. Since events with shower cores close to the border of the SD array might trigger enough stations to pass the T4, but these triggered stations

could represent just the outer parts of the footprint on the ground leading to a false energy and arrival direction assignment. Due to inactive stations in the SD array, the 6T5 is also applied to

For the highest energies where the flux is very small, this requirement is relaxed towards a 5T5 to increase the statistics enabling a more detailed study of arrival directions and thus, anisotropy studies. Simulations have shown that the hereby induced increase in error on the reconstructed shower properties is negligible.

In the end, one is interested in the efficiency of this trigger chain, i.e. at which energy all EAS are detected. Therefore, a Monte Carlo study has been performed with three different types of primary particles (protons, Fe-nuclei, photons) leading to a full acceptance of the array at an energy of $3 \cdot 10^{18}$ eV for the hadronic component of UHECRs (cf. Fig. 4.8) which was confirmed by cross-checks on real data (see [59]).

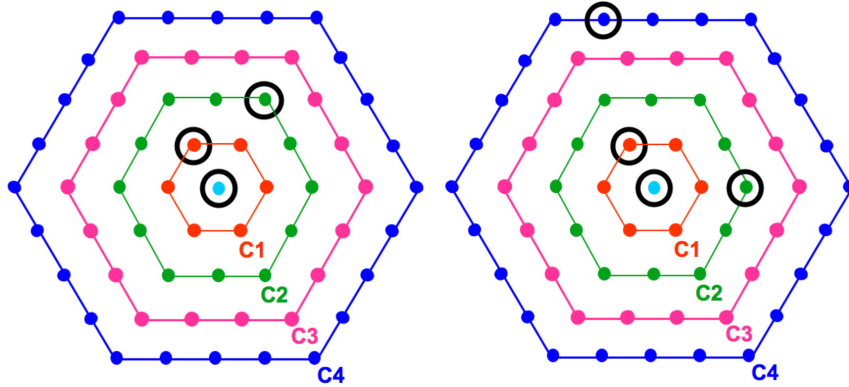


Figure 4.7: The two allowed T3 configurations. *Left Panel:* The ToT2C₁ & 3C₂ configuration including solely tanks with a ToT-T2. *Right Panel:* The 2C₁& 3C₂ & 4C₄ configuration in the case of at least one TH-T2 trigger within the event structure, (taken from [59]).

4.2.2 Reconstruction and Resolution

We discuss in the following the reconstruction procedure using the same event as in the section regarding the FD. It is a golden hybrid, i.e. FD as well as SD could reconstruct the event separately. This event has passed the complete trigger chain being a 6T5 which can be seen in Fig. 4.9. While the area of the filled circles is proportional to the signal strength in a station, the colour indicates the time of detection in one station (yellow early, red late). In addition, grey stations represent accidentals having an improper time information.

The reconstruction starts with an estimation of the shower core by calculating the barycentre of all stations containing a signal. According to the time information a planar shower front is fitted resulting in a preliminary arrival direction of the shower. Subsequently, the shower axis can be reconstructed and the signals are transferred to the shower plane perpendicular to the axis containing the shower core. In the following, all values, namely the distance to the shower core and the time of the corresponding station, are referred to with respect to the shower plane.

Starting with this first result, an iterative process is carried out to find the best

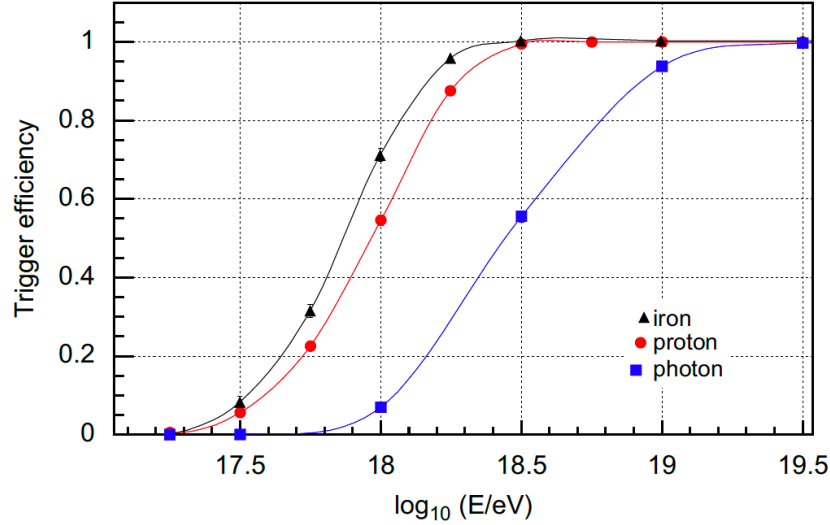


Figure 4.8: SD trigger efficiency as a function of Monte Carlo energy E for proton (circles), iron (triangles) and photon primaries (squares) and zenith angles integrated up to 60° . Lines are drawn to guide the eyes (taken from [59]).

estimate for the shower core, the agreement between the lateral shower profile and the fitted NKG-function as well as the arrival direction. The first step is the fitting of the lateral distribution function (LDF) to the data points around the shower core with a modified NKG-function:

$$S(r) = S(1000) \cdot \left(\frac{r}{1000 \text{ m}} \right)^{-\beta} \left(\frac{r + 700 \text{ m}}{1700 \text{ m}} \right)^{-\beta}, \quad (4.1)$$

with r being the distance to the shower core in meters and $S(1000)$ the extrapolated signal strength at 1000 m distance. $\beta(\theta) = 0.9 \sec(\theta) - 3.3$ is the slope of the lateral distribution function, a fixed parameter, and θ the reconstructed zenith angle of the air shower. The application to the example event can be seen in Fig. 4.10.

As proposed amongst others by Stanev [16], the value of $S(r)$ at a fixed distance instead of the complete curve is used to estimate the energy of the primary particle. It has been shown in [60], that for the PAO 1000 m from the shower core is optimal. Therefore, $S(1000)$ given in VEM is the first step towards an estimation of the energy.

Now, the challenge is to relate this variable to the direct energy measurement of FD. Here, one has to keep in mind that the more a shower is inclined the more it is attenuated due the longer way through the atmosphere. This leads to a smaller signal in a station in the end considering the same energy of the primary particle. From the presumption that no anisotropy in the arrival directions in local coordinates is present, one can conclude that the distribution of the flux J should be constant:

$$\frac{d}{d\phi} J = 0 \text{ and}$$

$$\frac{d}{d(\cos(\theta))} J = \frac{d}{d\theta} \left(\frac{1}{A_{eff}} \frac{d^3 N}{dt dE d\Omega} \right) = 0,$$

with θ and ϕ the zenith and azimuth angle respectively, $A_{eff} = A \cdot \cos(\theta)$ the effective area of the SD array projected on the shower plane and $d\Omega = d \cos(\theta) d\phi$ the solid angle. Using the second equation and focusing on the θ -dependency as well as on the energy, the following equation is obtained:

$$\frac{dN}{dEd(\cos^2(\theta))} = \text{const..}$$

By integrating over a certain energy range one can see that the shape of $dN/d(\cos^2(\theta))$ is independent of the energy, namely constant. This simplifies the procedure since only one curve has to be considered for all energies. In the end, one obtains a polynomial of second order in $\cos^2 \theta$ describing the progression of $S(1000)$. This function allows to convert the signal strength from one zenith angle to another. As a reference, $S(1000)$ at $\theta = 38^\circ$ is chosen, S_{38° .

In general, this approach is called the constant intensity cut method with the corresponding function $CIC(\theta)$ describing the behaviour depending on the zenith angle. From that we can conclude that $S_{38^\circ} = S(1000)/CIC(\theta)$.

S_{38° can now be correlated with the direct energy measurement of FD of golden hybrid events. Fig. 4.11 shows the correlation of the energy determined by FD in eV and the S_{38° in VEM. Note that the axes are log scaled. In addition, a straight line is fitted by means of a χ^2 minimization correlating both variables. The area below the dashed line is excluded to avoid a possible bias due to trigger efficiency effects. In the end an energy resolution of 5 % at 10^{20} eV could be achieved whereupon the 22 % systematic error is also transferred to the energy determined by the SD [61].

After the best values are found for the LDF, a more realistic model for the shower front is applied. The best description is achieved by using a parabolic shape (cf. Fig. 4.12) leading to a more accurate reconstruction of the arrival direction and thus, to a small change in shower axis and plane. Regarding the uncertainty of the angular reconstruction, an event-on-event based estimation is chosen called time variance model [62]. Assuming a Poissonian distribution of the particles in the shower front and accounting for the uncertainty of the time measurement, the variance of timing of the shower can be calculated which propagates directly into an uncertainty of ϕ and θ , the angles in the local coordinate system. For energies larger than 10^{19} eV and number of triggered stations larger than six, a resolution of $\sigma < 1.0^\circ$ is obtained. Studies using golden hybrids obtain comparable values [56]. The fitting of LDF and timing structures is iterated several times until it converges.

4.3 Enhancements

Due to the good performance of the southern site of the Pierre Auger Observatory, it offers the unique opportunity to use an excellent environment as a reference to approach further open questions and extending hereby the scientific purpose. Particularly interesting is the transition region between the second knee and the ankle where it is assumed that galactic sources die out and an extragalactic origin is held responsible for UHECRs. Two correlated questions are connected with this feature:

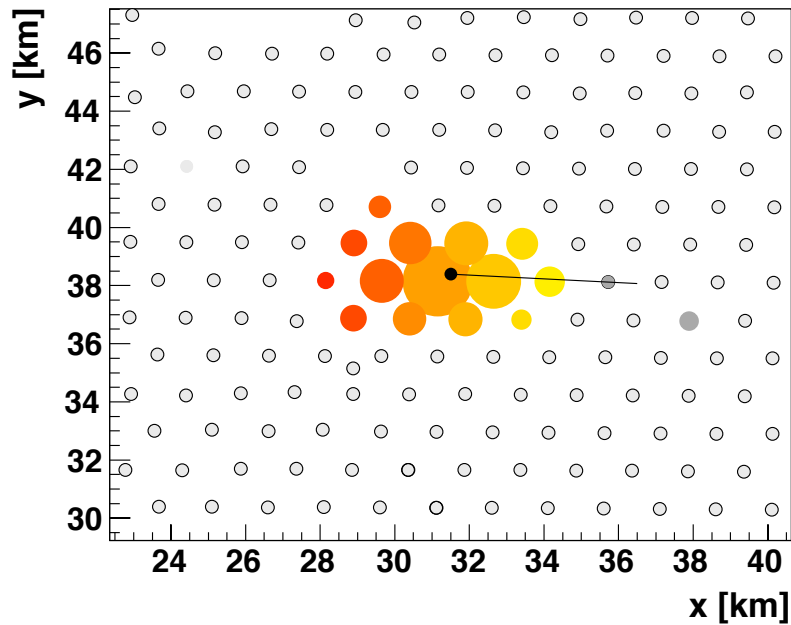


Figure 4.9: An example surface detector event fulfilling the 6T5. Each station is symbolized by empty circles. If stations are triggered they are replaced by filled circles with their area corresponding to the strength of the signal in the station. The colour represents the timing of trigger (yellow early, red late). Grey stations are too far away from another triggered station or have an time information which is not consistent with the reconstructed lateral distribution. Both types are flagged as accidental. The black line marks the reconstructed arrival direction.

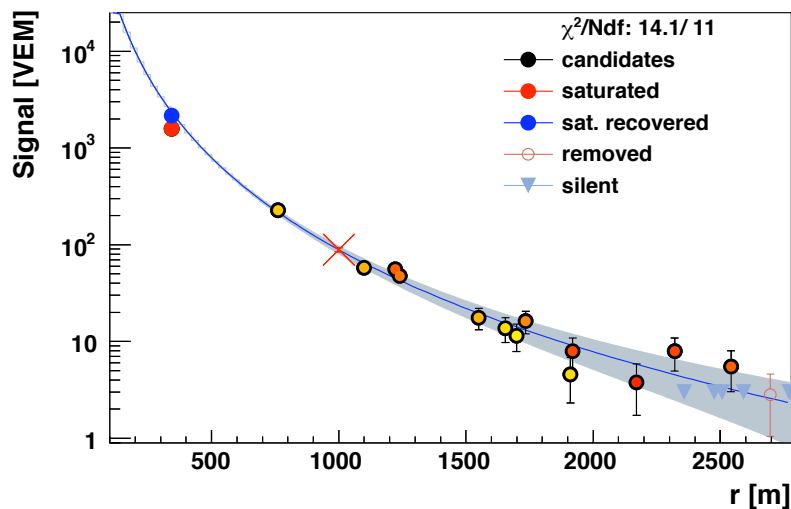


Figure 4.10: In blue one can see the final fit of the modified NKG function to the data points colour-coded corresponding to their timing. Grey triangles mark stations having detected no signal but according to the function could have seen one. The red point represents a saturated station whose true signal could be recovered (blue circle).

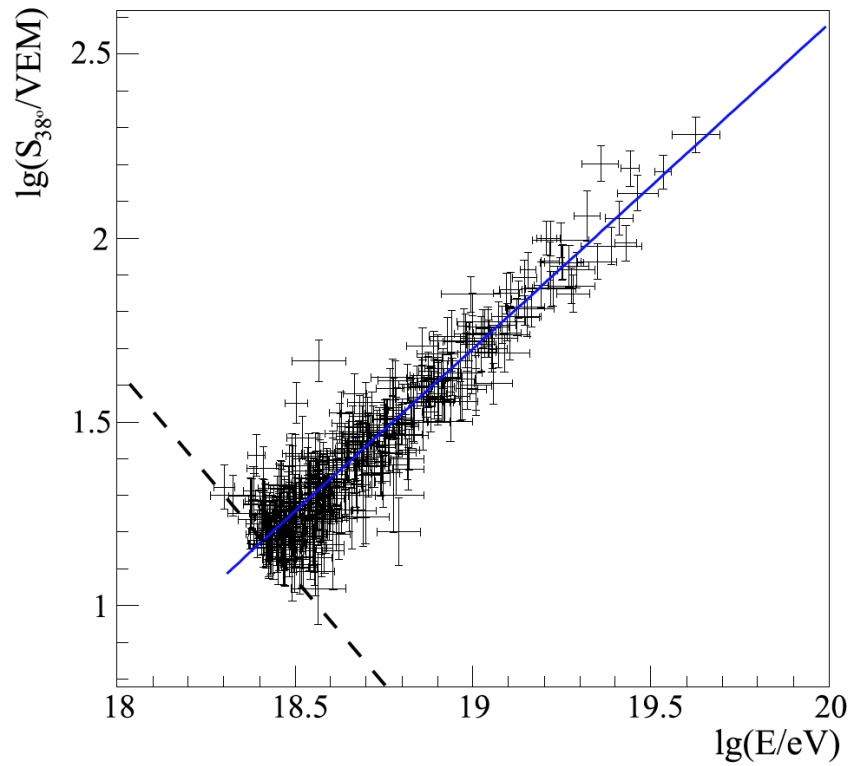


Figure 4.11: Correlation between $\lg E_{FD}$ and $\lg S_{38^\circ}$ for the 387 hybrid events used in the fit. The solid line is the best fit to the data. Events below the dashed line were not included in the fit.[61]

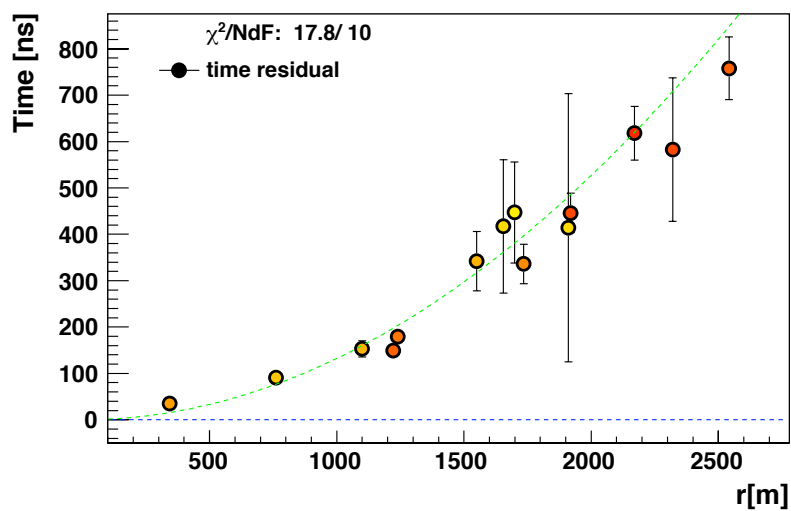


Figure 4.12: Displayed is the final timing fit of the event using a parabolic function. The colour code is the same as in Fig. 4.9 (yellow represents early, red late).

1. What is the distribution of X_{max} in that region depending on the energy?
2. How large is the difference between the estimated number of muons in air shower simulations compared to data?

Both questions tend to determine the composition of the primary particles which in turn can possibly help to resolve the puzzle of the origin of cosmic rays. The latter question is motivated by KASCADE-Grande results where the direct measurement of the muonic component of EAS was compared to the predictions of simulations [63].

With the original design of the PAO these questions cannot be answered due to the inability of the surface detectors to distinguish between muons and the electromagnetic component as well as the existing lower energy threshold of $3 \cdot 10^{18}$ eV where the full efficiency is reached. Therefore, two extensions called AMIGA and HEAT, which are described in the following, were designed to solve these issues.

In addition, it was decided to test the technique measuring the coherent electric field produced by charged secondary particles deflected in the magnetic field of the Earth by means of a large scale radio detection array (AERA). Simulations show that all shower parameters of physical interest can be reconstructed from the data [64].

By placing all new experiments in the vicinity to the Coihueco site (see Fig. 4.13) they benefit from each other and from the regular array following the idea of cross-calibration.

However, plans exist to test even more new techniques and enhance the observatory with further detectors, but we discuss here only those which have matured the most over the recent years.

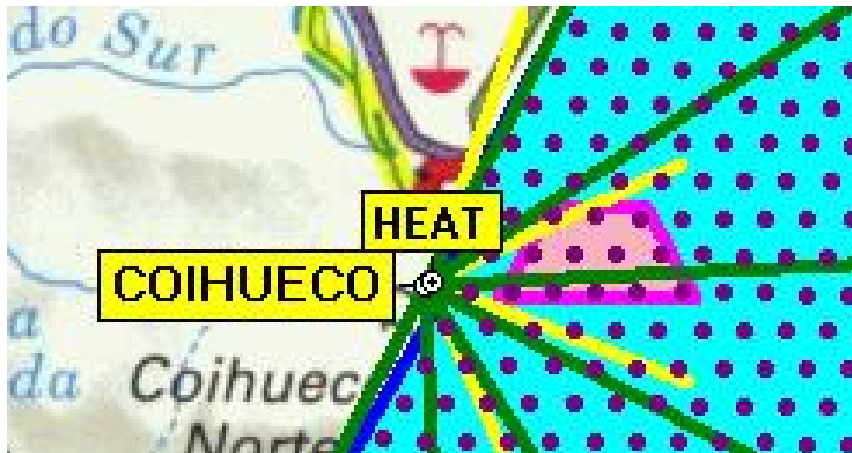


Figure 4.13: Pictured is the north-western border of the PAO where all the enhancements are located. HEAT close to Coihueco and the pink shaded area covers AERA and AMIGA.

4.3.1 High Elevation Auger Telescopes (HEAT)

The HEAT enhancement consists of three additional fluorescence telescopes housed separately in inclinable bays located roughly 250 m away from the Coihueco telescope

building (see Fig. 4.14). By operating in the non-tilted mode, they have the same field of view above the horizon as a standard FD telescope and HEAT is counted therefore as the 5th eye. However, its actual purpose is to extend the lower energy bound to $10^{17.5}$ eV. By looking higher into the atmosphere ($30^\circ - 58^\circ$), a large amount of the longitudinal shower profile of an EAS with a smaller energy can be seen due to its earlier development in the atmosphere. Thus, X_{max} is more often in the field of view of HEAT than in the field of view of the standard FD. This allows to study the composition at these energies. Nevertheless, the ability to operate in two modes has not only the advantage of cross-calibration between HEAT and Coihueco but also to reconstruct those showers which would otherwise lead to too large uncertainties on the observables of interest, e.g. X_{max} . An example for such an event is shown in the bottom panel of Fig. 4.15, without HEAT this event would have led to an unreasonable reconstruction.

The first HEAT events were recorded in the beginning of 2009 already in coincidence with the standard FD telescope of Coihueco (see top panel of Fig. 4.15). By now, it is working on a regular basis integrated into the normal data acquisition system. Future plans regard an extension of the SD array in the direction of HEAT to have more hybrid events in the data set: the EAS seen by HEAT operating in inclined mode are closer to the telescopes due to their lower energy. Without the extension of SD they might not trigger a station which would lead to a significant increase of the uncertainty in the reconstruction. Particularly at the lowest energies, this is counter-productive considering the motivation of HEAT.



Figure 4.14: Photo of all three inclined HEAT telescopes. Bay 1 to Bay 3 from left to right.

4.3.1.1 TILT-Monitoring

Due to the dependence of the event reconstruction on the optical stability of the camera-mirror system it is necessary to monitor the relative changes in inclinations

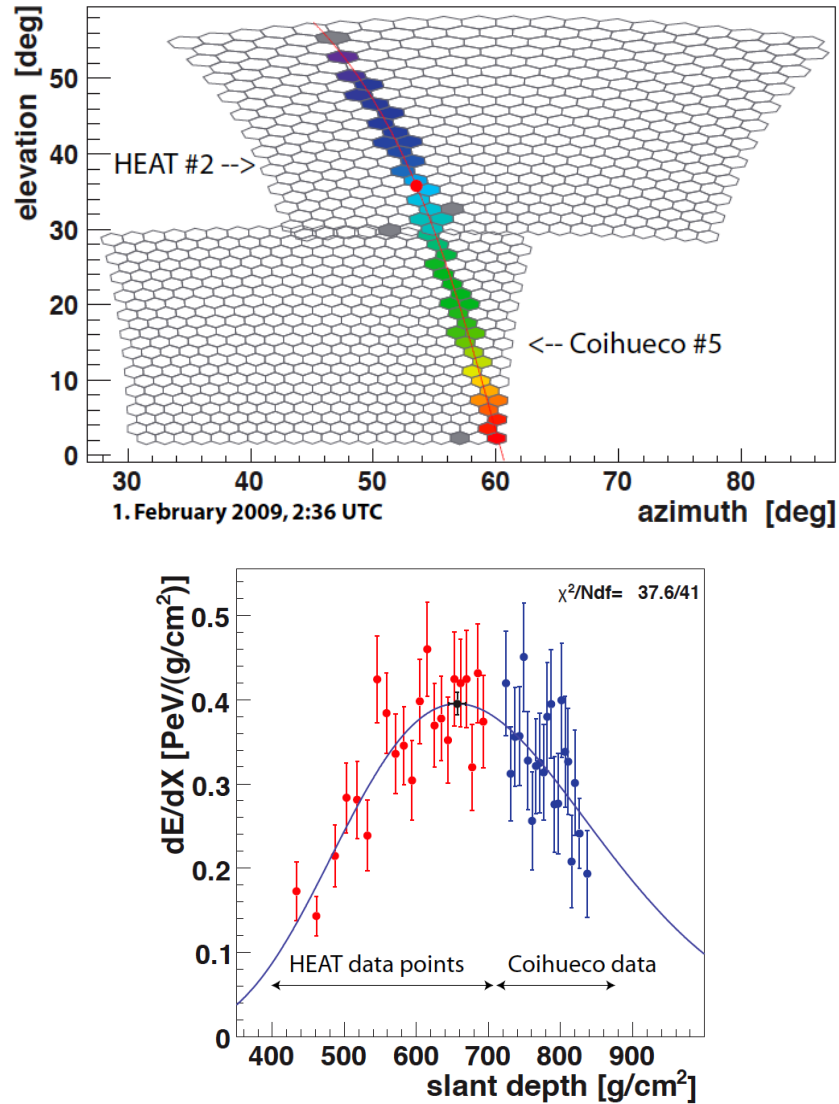


Figure 4.15: A HEAT and Coihueco Stereo Event. *Top Panel:* Colour-coded is the time structure of this event from blue to red which corresponds to several μs . The red thin line represents the fit of the shower detector plane. *Bottom Panel:* Measured longitudinal profile including a Gaisser-Hillas fit. Only the usage of both, HEAT (red, left) and Coihueco (blue, right) gives a reasonable value for X_{max} (both taken from [65]).

and distances of the telescope components. In this way, a stable performance during periods of data taking independent of the operation mode can be ensured. For this purpose, we introduce here the Aachen TILT-Monitoring System consisting of four inclination and four distance sensors. They are mounted at important reference points (cf. Fig. 4.16) of the system measuring relative changes during operation. A first prototype of this system was developed by J. Calvo de No (see [66]) during his diploma thesis in collaboration with M. Leuthold. His thesis includes the selection process for the sensors, the corresponding calibration and the results of a test run of the prototype installed in Bay 1 of Los Leones.

For the final installation the two additional set-ups had to be calibrated and changes on the hard- and software were necessary.

The configuration for all HEAT bays can be found in Fig. 4.16. The description is made from the point of view of someone standing behind the mirror looking in the direction of the shutter. For the inclination sensors:

1. base of the camera
2. top of the camera,
3. top middle of the mirror and
4. left side of the shutter.

The distance sensors are fixed at the first mentioned position:

1. top left of mirror to top left of the camera
2. top right of mirror to top right of the camera
3. bottom left of the mirror to the shutter and
4. center of the mirror to a point beneath the camera.

In this way, it can be monitored for example whether the distance and/or the relative inclination between the camera and the mirror are changing. If the positions are differing from what is expected by the reconstruction, this would lead to a false event reconstruction. In addition a similar study can be performed including the shelter of the telescope. Outside influences like strong winds or the frequent tilting could deform the shelter and thus, cause systematic errors on the reconstruction.

Sensors and Calibration

In the following, we briefly describe the incorporated sensors and focus mainly on the applied calibration procedure.

Inclination

The basic requirement for this type of sensor is a good resolution (design goal: $\Delta\alpha \approx 0.1^\circ$) between the down- and the up-mode. Thus, a minimum range of 30° has to be covered. In the end the the following sensor was chosen:

- Dual Axis Inclinometer (**DAS-15-MC-RS232**), see left panel of Fig. 4.17.

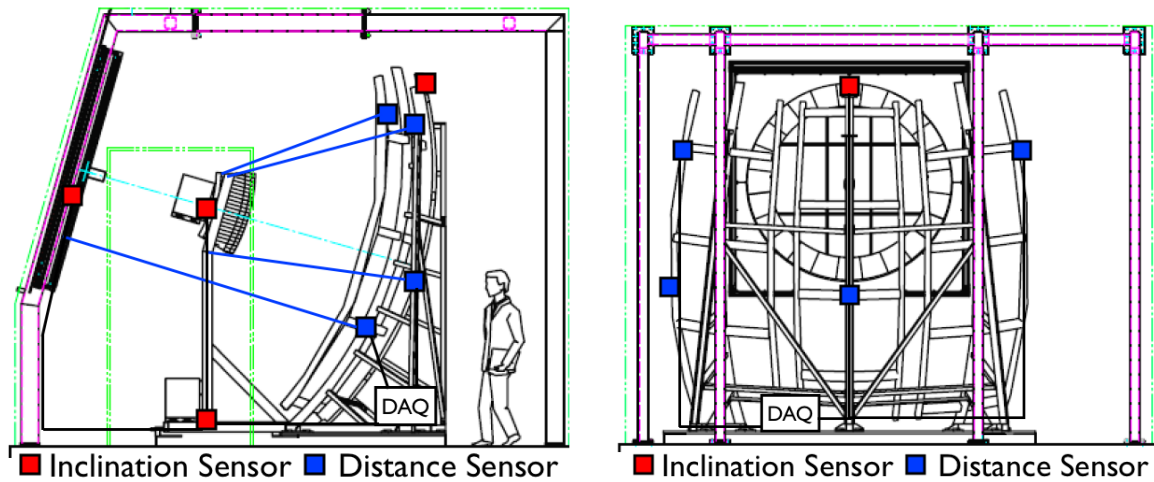


Figure 4.16: Position of the sensors within one bay: seen from the side (left) and from the back (right). In red the inclination and in blue the distance sensors.

with a measurement range of $\pm 15^\circ$ in x and y direction. This is sufficient because they are mounted with a preset inclination of roughly 15° .

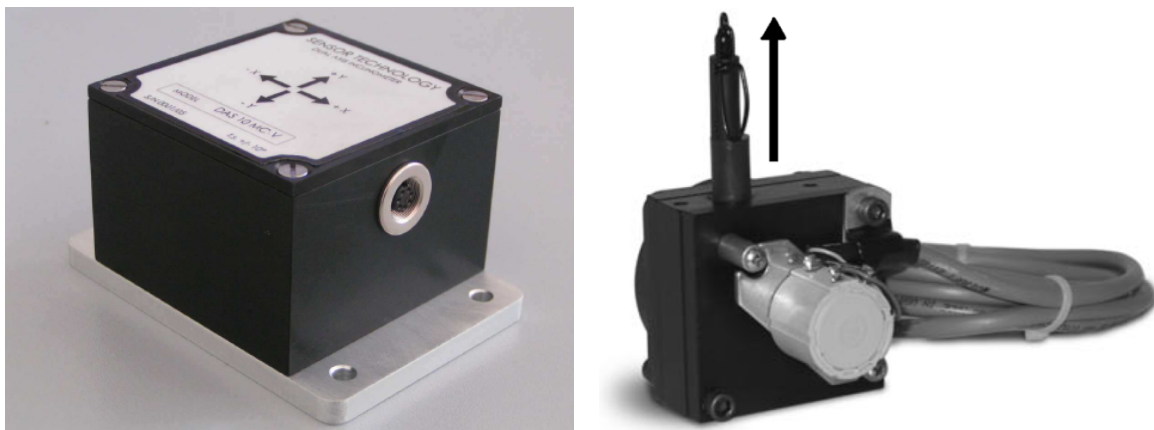


Figure 4.17: Installed sensors in the HEAT shelters. *Left Panel:* Inclinometer Sensor DAS-15-MC-RS232. *Right Panel:* Distance sensor PTX101 (Potentiometer).

We demonstrate the procedure on one inclination sensor with the serial number 143, since it is similar for all others. For the calibration we used a goniometer with a range of $\pm 10^\circ$. Thus, a continuous measurement over the complete range was impossible. Therefore, an angle piece of 10° was included in the set-up (see Fig. 4.18) leading to three measurement-intervals:

-17°	-	-3°	(w/ angle piece)
-7°	-	7°	(w/o angle piece)
3°	-	17°	(w/ angle piece)

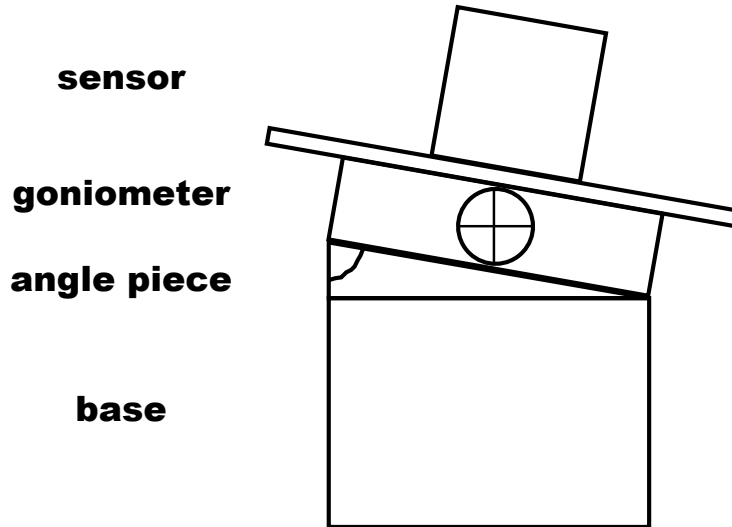


Figure 4.18: The set-up for the calibration of the inclination sensors. It consists of base, goniometer and sensor. For measurements with large inclinations an angle piece is included.

By making these three measurements overlap it was possible to account for constant offsets caused by the modification of the set-up during the calibration. The stepsize is $0.082^\circ \pm 0.008^\circ$ given by the accuracy of regulating screw.

For each adjustment three measurements are made and the mean is calculated. In Fig. 4.19 one can see the the difference of measurement and true value against measured inclinations including the anticipated small shifts mentioned above between the individual measurements.

Even without correcting the discontinuities, one can see that the deviation from a linear behaviour is 0.2° at the maximum. However, the requirement is a factor two smaller. Therefore, we corrected for the systematic shifts by calculating the mean distance between the overlap regions and adjust the data accordingly (cf. Fig. 4.19). At last, we fit the data points with a polynomial of seventh order. The corresponding residuum and projection on the y-axis can be found in Fig. 4.20 including a Gaussian fit giving a $\sigma < 0.01^\circ$. Thus, the resolution is now well below the required value. For the other sensors the same resolution was obtained by using a similar procedure, i.e. for some a polynomial of fifth degree was sufficient. The calibration constants a_i for all sensors can be found in Tab. B.1 in the appendix. The coefficient a_0 is omitted since we are interested only in relative changes.

In addition, also the the y-axis was calibrated over the range: -10° to 10° . Hence, we could directly fit a polynomial of seventh order leading to a similar resolution smaller than 0.01° . The corresponding constants can be found in Tab. B.2.

Distance

Similar to the inclination sensors, the selection procedure was performed by J. Calvo

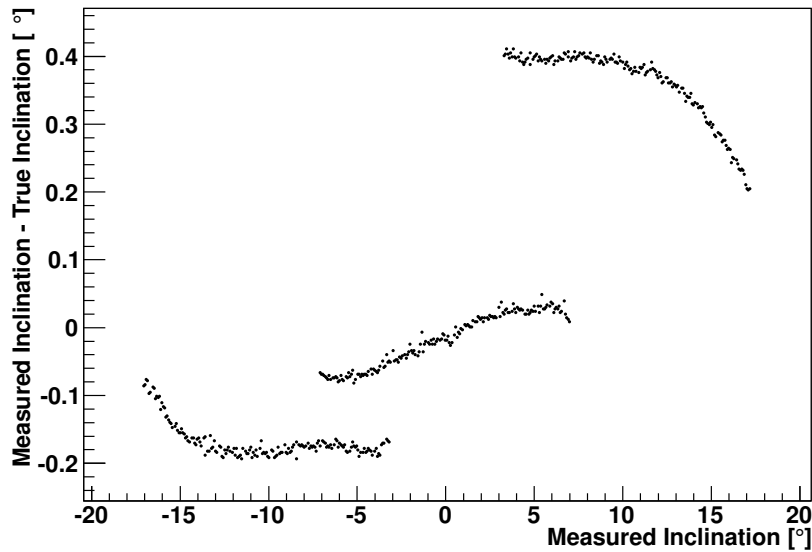


Figure 4.19: Displayed is the difference between measurement (given by the inclination sensor) and true value (given by the goniometer) against the measurement of the inclination sensor.

de No and is presented in [66]. In the end a potentiometer (see Fig. 4.17) with a maximal elongation of 5 cm was chosen to provide the needed accuracy. However, the distances to be monitored are much longer and therefore, temperature stable cables are attached to the sensors enlarging the total distance without losing the precision on relative changes. The calibration of the two additional distance sensor sets was carried out by N. Scharf and will be part of his Ph.D. thesis.

4.3.1.2 TILT-Setup

Software

For the long term test of the TILT set-up a basic version of the program was written in LabView by J. Calvo de No. This old version directly sent the data via email to him from the laptop installed in Los Leones. As a result of changed requirements, some modifications became necessary. The data is now filled every second in the general monitoring database of the PAO by means of an ODBC connection. In addition, some minor changes affect the interactive plotting and the start up of the program.

A typical cycle goes as follows:

1. computer boots
2. after start up the program runs automatically
3. trying to establish connection to sensors (repeats until success)
4. reading data and filling into database (every second).

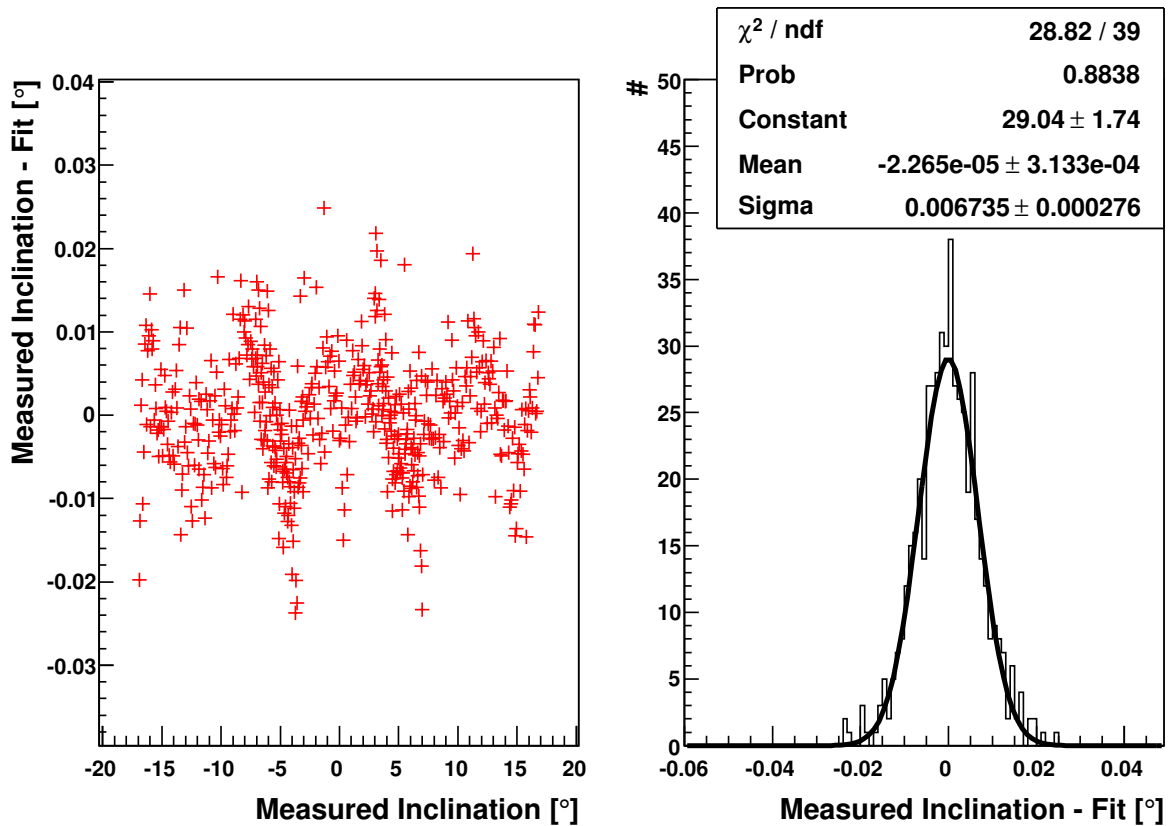


Figure 4.20: *Left Panel:* Residuum for the polynomial fit. *Right Panel:* Projection of the Residuum on y-Axis. The Gaussian fits leads to an accuracy of 0.006° .

Hardware

The concept (see Fig. 4.21) of the current installation is basically the same as for the prototype, just the weather sensors are removed from the set-up and the cable extending the measurements range of the distance sensors are changed from coated carbon fibres to temperature stable cables due to failures during the long term test. In the following we describe the installation of the hardware. In the center of the set-up is the embedded PC housed in a metal box placed behind the mirror. It is connected via an USB-cable on the one hand to an RS232 to USB-hub, which in turn is connected by four cables through a power supply box to the inclination sensors. On the other hand the PC is connected to the μ -Box which in turn is connected to the four distance sensors. The μ -Box is provided by the producer of the sensors digitizing the analogue signals of the potentiometers. Additional power supplies and data acquisition parts (e.g. the μ -box) are also placed within the metal box.

In 2010, the installation process was completed and since then the system is running stable apart from general power failures. As soon as the power is turned on again, the PC reboots automatically and starts data taking. For an analysis of the first data we refer to the Ph.D. thesis of N. Scharf (to be completed in 2011).

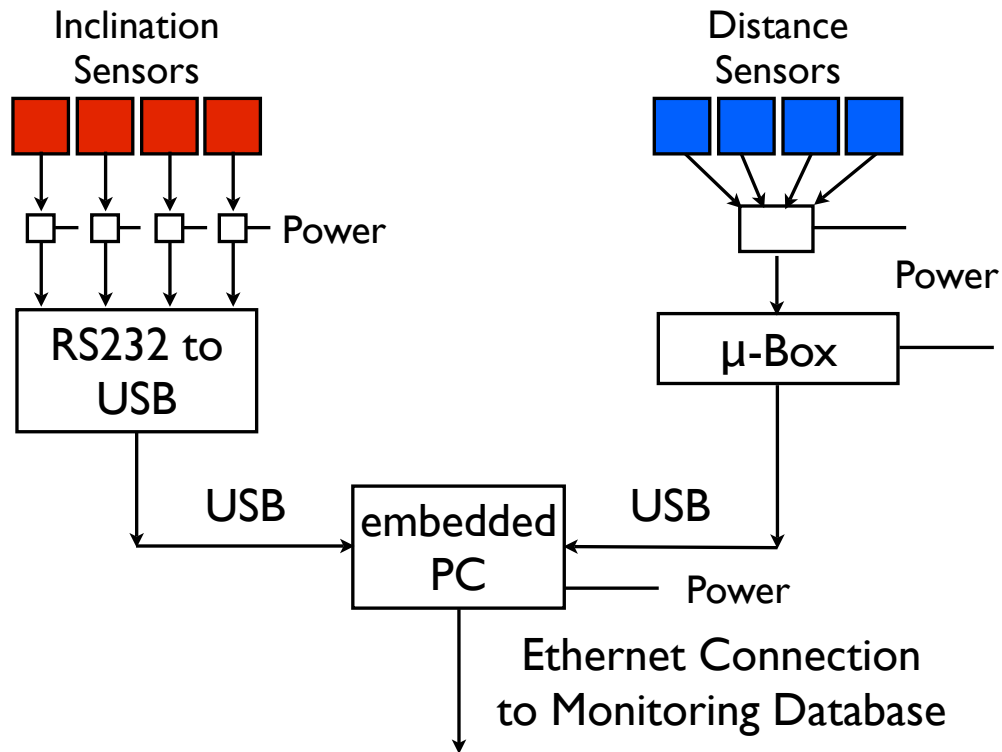


Figure 4.21: Scheme of the current TILT-Setup

4.3.2 Auger Muons and Infill for the Ground Array (AMIGA)

AMIGA consists of two parts. The first is the Infill, an additional hexagonal grid of 85 SD stations with smaller spacings of 430 m and 750 m respectively covering an area of 23 km². It is embedded in the regular SD array. The decrease in distance among the stations lowers the energy threshold of a T3 since EAS with $E \approx 10^{17}$ eV have a significantly smaller footprint.

Muon counters form the second part. Each one of the stations of the Infill has its own muon counter partner buried circa 2 m below the ground to suppress the electromagnetic component of the shower. A sketch for one station is shown in Fig. 4.22. One counter consists of 4 modules which in turn have 64 scintillator strips each attached to the PMT in the center of the corresponding module. A unitary cell consisting of seven AMIGA detector pairs, SD station and muon counter, is under construction at the moment.

Although, it is assumed that the electromagnetic component is negligible an additional detector type is constructed called BATATA. Three layers of plastic scintillators are buried to study the electromagnetic contamination depending on depth. It will be placed in the center of the AMIGA site surrounded by three SD stations on the surface. This will test the results obtained from simulations for the optimal depth for the muon counters of AMIGA.

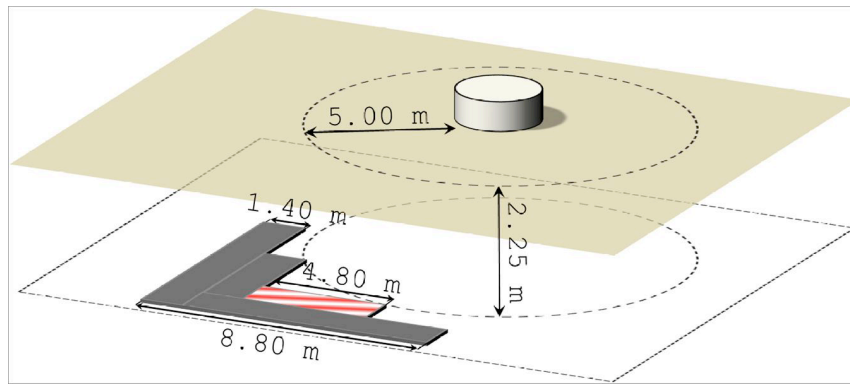


Figure 4.22: The AMIGA layout for a typical Infill station and corresponding Muon Counter. [65]

4.3.3 Auger Engineering Radio Array (AERA)

The Auger Engineering Radio Array is designed to study EAS caused by UHECRs by measuring the coherent radio emission of secondary shower particles deflected in the Earth's magnetic field. Its aim is to extend the operative range of this technique to energies close to $10^{18.7}$ eV, which is approximately one order of magnitude higher than current dedicated radio experiments, e.g. LOPES. After a good understanding of processes generating this kind of emission is achieved, it will be explored which variables of EAS are accessible and to what precision. At this, it will be supported due to its location by the PAO and the above described extensions.

The current design of the array covers an area of 20 km^2 and is compromised of three construction stages. The first one consisting of 24 stations visible as the dense inner part in Fig. 4.23 with a spacing of 150 m on a triangular grid was completed in 2009 and is taking data at the moment (spring 2011). In a second and third stage it will be extended with 50 and 60 further stations respectively whereupon the spacing increases to 250 m and 350 m.

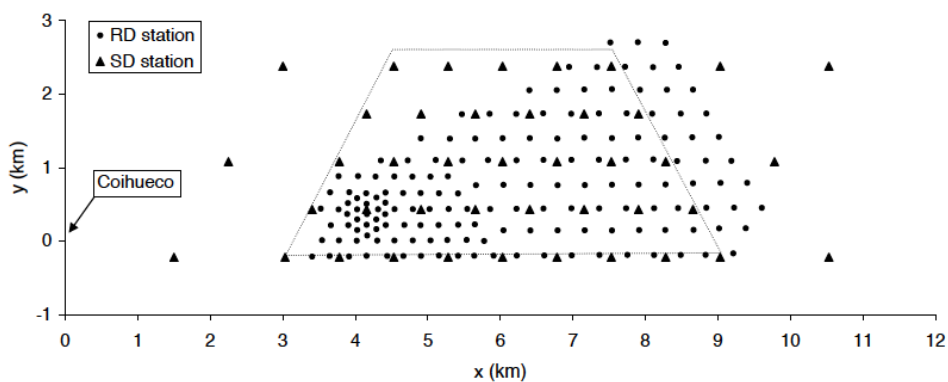


Figure 4.23: The current design of the grid for the radio antennas (circles) of AERA embedded in the regular SD array (triangles). The coordinate system is aligned with respect to Coihueco at the (0 km, 0 km) position. (taken from [65])

5. Autocorrelation Methods

In chapter 3, a search for anisotropy in the arrival directions of the highest energetic events measured by the Pierre Auger Observatory has been presented. It is based on the correlation of UHECRs with AGN of the Veron-Cetty and Veron catalogue within an angular scale of 3° . Thus, this study depends strongly on the choice of the catalogue and hence, on the contained astronomical objects. It can be argued that there is no connection between UHECRs and AGN, although they are prominent source candidates (see chapter 3.3.1.1). Furthermore, many source models exist beside AGN claiming to be able to explain the origin of UHECRs. To avoid the need to make assumptions about sources or their distribution, the concept of *Autocorrelation* can be adapted, i.e. a data set is tested whether it is cross-correlated with itself or not. In the context of the recently obtained positive anisotropy signal, new algorithms were developed such as the 2pt+ correlation function [67] and those discussed in this chapter to perceive an intrinsic anisotropy.

In the following, three methods are introduced starting with a representation of the *2pt-Correlation Function*, followed by a *Minimum Spanning Tree* and a *Cluster Algorithm*. Since there are various kinds of possible anisotropy scenarios in nature, each of these methods is designed to be sensitive to a different kind, e.g. the Cluster Algorithm to small scale structures. An important requirement for all of them is the unbiasedness when applied to an isotropic data set, i.e. the probability distribution that isotropy agrees with isotropy for many test sets should be uniform between 0 and 1.

For all these methods, the data is used in equatorial coordinates (right ascension α , declination δ) unless it is noted otherwise. α and δ are the celestial equivalents to the terrestrial longitude and latitude, whereas the declination is zero at the equator, 90° at the north and -90° at the south pole in contrast to the standard mathematical spherical coordinates. A pair of these two angles defines an event A .

To explain exemplarily how a particular method works, we use a test data set consisting of 100 events based on an anisotropic probability density distribution displayed in Fig. 5.1. The estimator is computed according to the chosen method and a probability is assigned. To this end, the estimator is compared to the results obtained by analysing $n_{sets} = 10,000$ isotropic sets which represent the Null-hypothesis (H_0). They are simulated on the whole sphere reducing the probability density to a pure geometric one, i.e. in right ascension the density is flat and in declination follows a cosine behaviour. Considering the field of view of the PAO in the next chapter, a more complicated formula is introduced.

For all methods the same procedure is chosen. We use the reference distribution of the estimator based on H_0 to calculate the probability P by integrating over

that part of the histogram providing values deviating even more from the isotropic expectation than the test data set (see Fig. 5.1). The resulting value, which is normalized to the integral of the histogram, gives the probability that the test data set is based on the same scenario.

For these and following analyses we have used the *ROOT* framework [68].

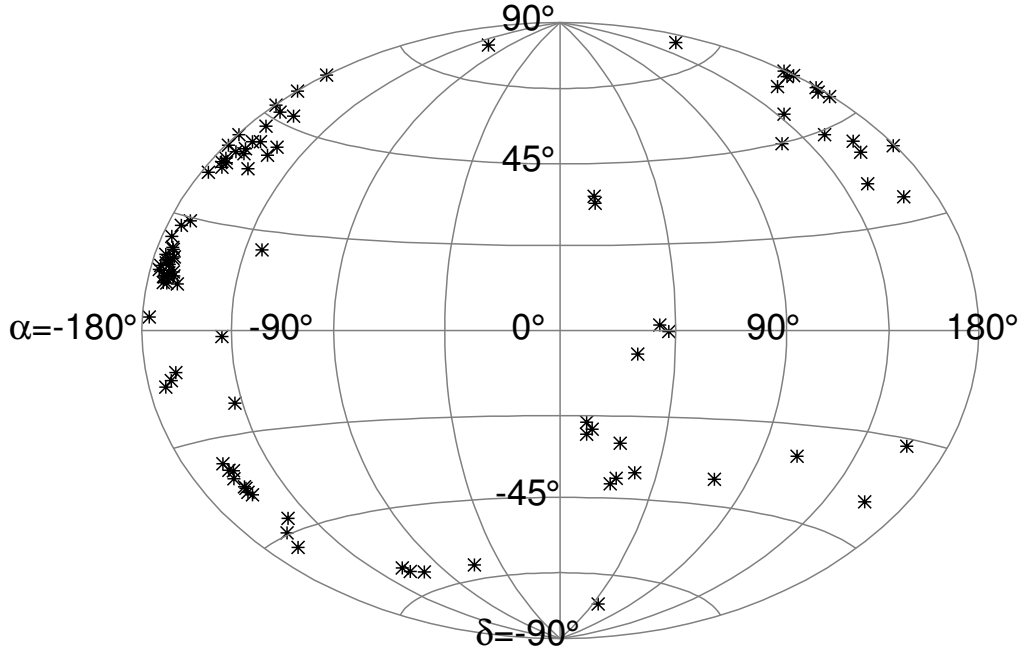


Figure 5.1: Example MC data set of 100 events based on the VCV catalogue (Hammer Projection).

5.1 2pt-Correlation-Function

The 2pt-Correlation Function is mostly referred to as the standard autocorrelation method. In many cases, the integral version is used, e.g. within the latest PAO publication regarding anisotropy studies[25] (cf. Fig. 3.7). In the differential representation, the angular distances between all events are calculated and filled into a histogram for instance. The integral version is then the resulting cumulative distribution based on this very histogram, i.e. each bin is equal to the integral of all bins below this one of the differential histogram.

We consider in the following the differential approach in two separated realisations using one angle (chapter 5.1.1, referred to as TwoPt) and two angles respectively (chapter 5.1.2, referred to TwoPtPlus) to test for a compatibility with the isotropic expectation.

5.1.1 Using one angle

We calculate the solid angle Ω between two events for all possible combinations. The result for the test data set (red,dashed) and the isotropic data sets (blue shaded and

mean in black) can be found in Fig. 5.2. Here, the difference is clearly visible, the angular distance between two events is enhanced on small scales for the anisotropic case (dashed, red). In addition, the probability density distribution for the isotropic realisations is shown in shades of blue using again 100 events per set. In this 2-dimensional histogram for a fixed bin i on the x-axis, the distribution along the y-axis is normalised to one.

As a measure for the discrepancy between the test set and the isotropic expectations, we use

- a negative log-likelihood (LL) and
- a Kolmogorov-Smirnov-Test (KS).

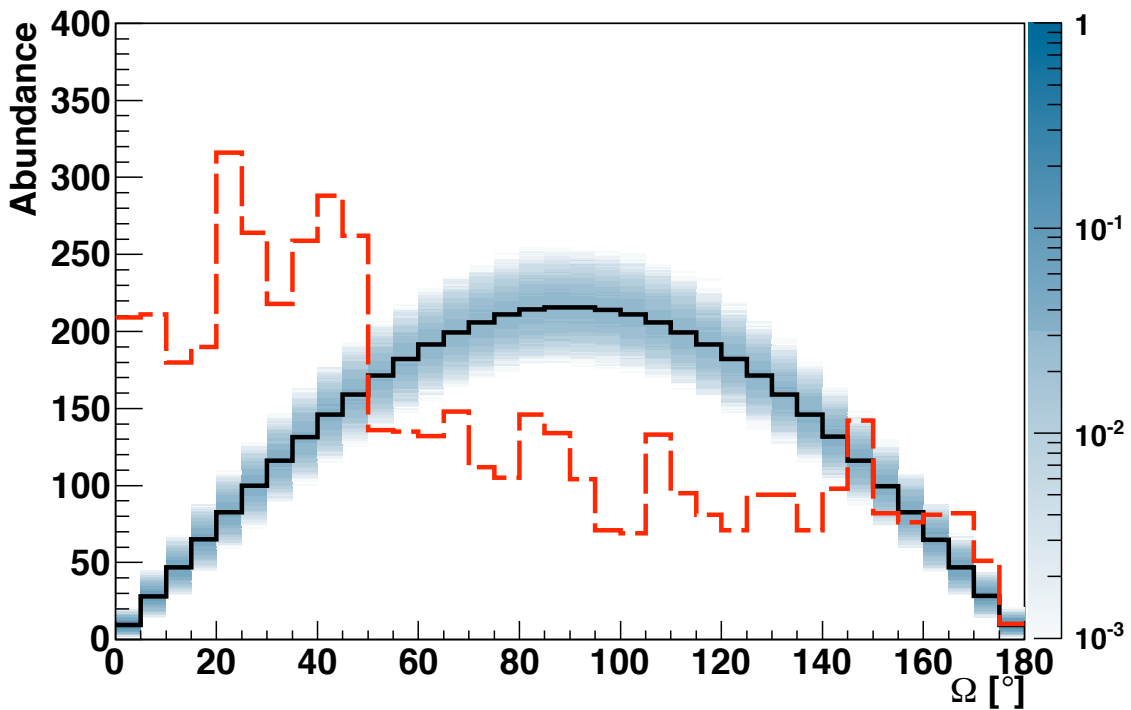


Figure 5.2: Solid Angle Distribution for the test data set is shown as dashed (red) line, while the isotropic density distribution is color coded in blue shade. It is based on 10,000 isotropic MC sets. The mean of all isotropic MC sets is displayed as a black line.

While the first one compares the distributions of the isotropic and the example MC sets over the complete range, the second focusses on the point where the distance between those two is maximal.

Loglikelihood-Test (LL)

For the loglikelihood test the standard formula is used:

$$LL_{data} = -2 \cdot \sum_{i=1}^{n_{bins}} \ln P(\Omega_{i,data} | \Omega_{i,iso}). \quad (5.1)$$

Hence, we sum the logarithm of the probability P that we measure $\Omega_{i,data}$ when $\Omega_{i,iso}$ is expected on average. In general, one would use the Poissonian errors for bins of a histogram, but here they are correlated leading to a deviation from this assumption. Therefore, we rely on the density distribution based on the 10,000 isotropic sets to estimate the probability in a bin (i, j) . By performing the same analysis for the isotropic as well as the test data sets, the method is reasonable.

As one can see from Fig. 5.2 many bins of the reference density distribution (in blue shade) are empty leading to infinite terms in equation (5.1). To avoid this, these bins are set to $\frac{1}{2 \cdot n_{sets}}$, the half of the minimal possible value.

By applying this method to n_{sets} isotropic sets we obtain a histogram for LL_{iso} (see left panel of Fig. 5.4). In addition, the result of the test data set LL_{test} is also included, visible as a red line. Subsequently, the probability is then calculated by integrating. The lower and upper integration limits are $LL_{test} = 634$ and infinity, respectively. In this very anisotropic example the probability that the test data set is based on the isotropic expectation is smaller than 10^{-4} .

Kolmogorov-Smirnov-Test (KS)

The main advantage of this kind of test is that it already shows a good performance on low statistic samples, i.e. the error of the first kind is small. However, it is known that the error of the second kind is enhanced leading to the acceptance of the Null hypothesis when it has to be rejected.

Compared to the LL-Test, the KS-Test considers indeed the complete histograms, but for the estimator calculation only one bin is crucial in the end. Furthermore, the mean distribution of the isotropic model is used instead of the probability density.

Here, the estimator is computed in the following way:

$$\delta_{KS,test} = \max |F_{test}(\Omega_i) - F_{iso}(\Omega_i)|, \quad (5.2)$$

with F the corresponding cumulative normalized distribution (see Fig. 5.3) based on the histograms (black and red line) shown in Fig. 5.2.

The index KS is added to avoid a confusion with the standard variable of the declination δ .

Within this method, the largest absolute distance between the mean isotropic distribution F_{iso} and the one for the test data set is calculated. As before, the larger the estimator the larger the deviation from isotropy. To assign a probability to the final value of $\delta_{KS,test} = 0.29$ for the anisotropic MC test sample, the value of δ_{KS} is computed for the 10,000 reference sets. The resulting histogram of $\delta_{KS,iso}$ and the result of the test data set can be seen in the right panel of Fig. 5.4. The agreement is calculated again by integrating over the resulting reference distribution from $\delta_{KS,test}$ to infinity giving a probability of $P < 10^{-4}$.

Combination

Both tests have their own assets and drawbacks, while the KS needs less statistics to be conclusive than the LL, LL is more precise given a deviation from the Null hypothesis is present. Therefore, we combine them in the sense that it is required for the a test data set to exceed LL_{test} as well as $\delta_{KS,test}$.

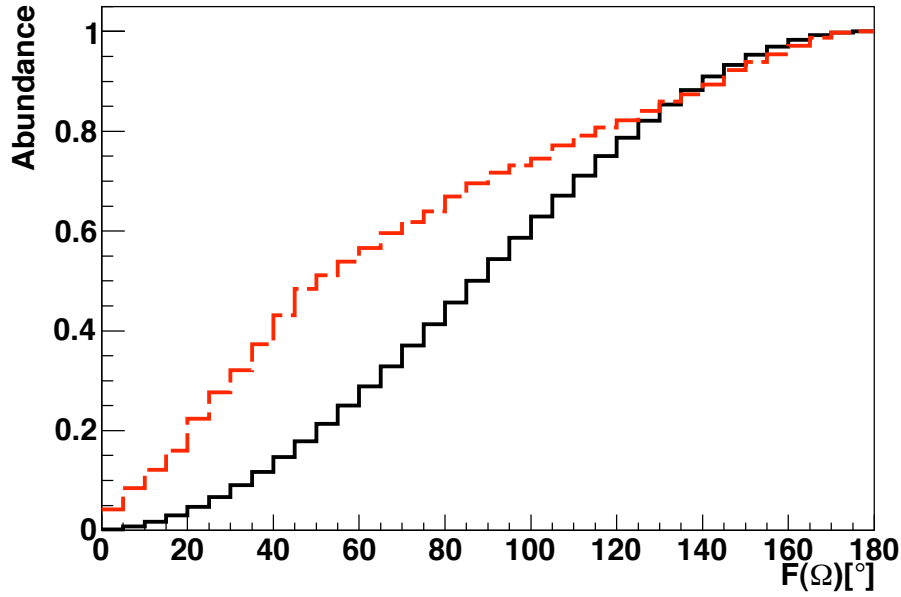


Figure 5.3: Cumulative angular distance distribution $F(\Omega)$ for isotropy (black) and the test event sample (red,dashed).

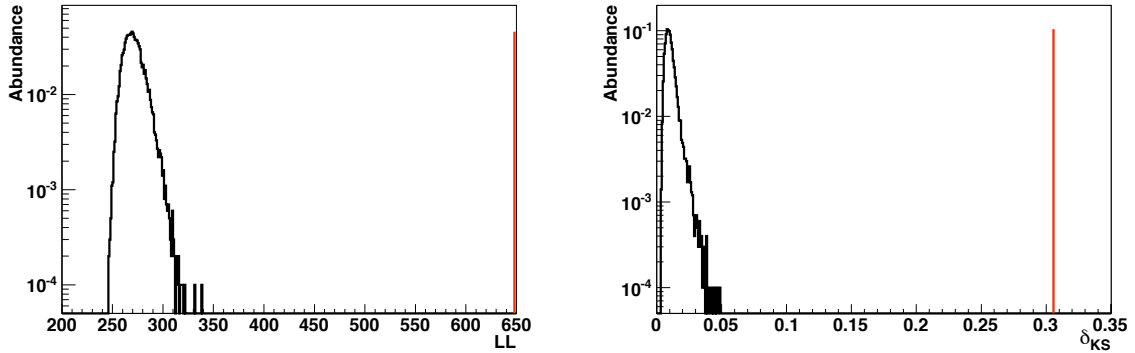


Figure 5.4: In both plots the estimator distribution of the isotropic expectation is shown (black) as well as the result for the test data set (red line). In both cases the reference and test data set are clearly distinguishable. *Left Panel:* Distribution of LL_{iso} . *Right Panel:* Distribution of $\delta_{KS,iso}$.

However, for the probability calculation an additional step has to be introduced. In contrast to the procedure in the 1-dimensional case where in principle one value of the estimator corresponds exactly to one probability, now several points in the KS-LL-plane can give the same probability (cf. Fig. 5.5). This would lead to a bias towards smaller values of the probability (see Fig. C.1). Therefore, the resulting probability has to be corrected which is discussed in detail in appendix C.1. In this context, a reference probability distribution is needed based on an independent sample of isotropic sets. To compute the true probability of a test data set, the fraction of isotropic sets is calculated which gives an even smaller probability than

$P_{test,bias}$.

In the end, one obtains also a probability of $< 10^{-4}$ that H_0 and the test data set agree.

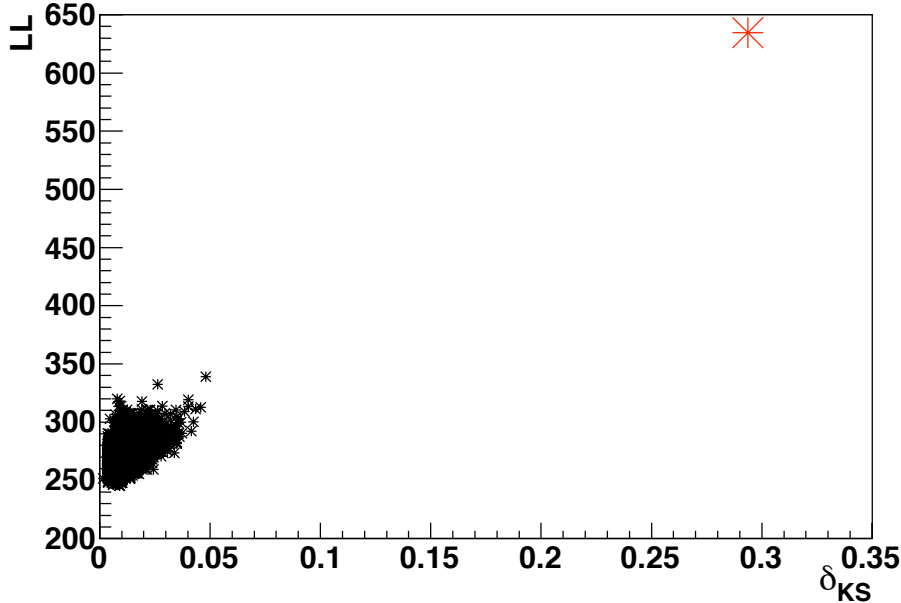


Figure 5.5: Estimator distribution (black small crosses) in the KS-LL plane based on 10,000 isotropic sets and the result of the test data set (red large cross).

5.1.2 Using two angles

The absolute angular distance between two events describes the relative position with respect to each other. However, the events are distributed on a 2-dimensional sphere, i.e. two independent coordinates are necessary to determine their absolute positions. Thus, more information can be gained by accessing this additional information. A disadvantage which is linked to this circumstance is the requirement to decide for a coordinate system. This was unnecessary if one is only interested in the relative differences.

Here, we restrict ourselves to the equatorial coordinate system (α, δ) , which can be justified by experimental arguments and is found among those mentioned in chapter 2. The ECS is the natural choice for the PAO since it is earthbound and in time independent in relation to one epoch. In addition, it has the advantage of yielding a flat distribution in right ascension allowing to recognize at an early stage deviations from this behaviour. In this way, the method is still independent of source assumptions since the focus lies solely on the detector and the associated conditions. The procedure is adapted from the version with only one angle (TwoPt), i.e. for each variable we use the KS- and LL-test and combine them to estimate the final probability. First, the angular differences in α and δ are calculated (cf. Fig. 5.6). It can be noticed that the $\Delta\alpha$ -distribution is flat for the isotropic case allowing to recognize easily a deviation from this constant behaviour. The reason for this lies

in the flat density distribution in azimuth (or here right ascension) on a sphere. In contrast to the behaviour of this variable, the density distribution of δ follows a cosine behaviour accounting for the smaller solid angles at the poles compared to those at the equator.

Applying both methods, KS and LL, to $\Delta\alpha$ and $\Delta\delta$ of the test data set the following results are obtained:

$$\begin{aligned} \delta_{KS}(\Delta\alpha) &= 0.33 &\implies P_{KS}(\Delta\alpha) &< 10^{-4} \\ \delta_{KS}(\Delta\delta) &= 0.078 &\implies P_{KS}(\Delta\delta) &= 1 \cdot 10^{-2} \\ \delta_{LL}(\Delta\alpha) &= 660 &\implies P_{LL}(\Delta\alpha) &< 10^{-4} \\ \delta_{LL}(\Delta\delta) &= 501 &\implies P_{LL}(\Delta\delta) &= 5 \cdot 10^{-4} \end{aligned}$$

The corresponding distributions can be found in Fig. 5.7. All methods indicate a deviation from isotropy only the KS test regarding $\Delta\delta$ is in the transition region close to 1% where the outcome is undecided.

Combing all variables and tests as in the previous method yields a final probability of $< 10^{-4}$ taking into account the needed correction of the probability (see appendix C.1).

A discussion regarding the comparison between TwoPt and TwoPtPlus can be found in the next chapter.

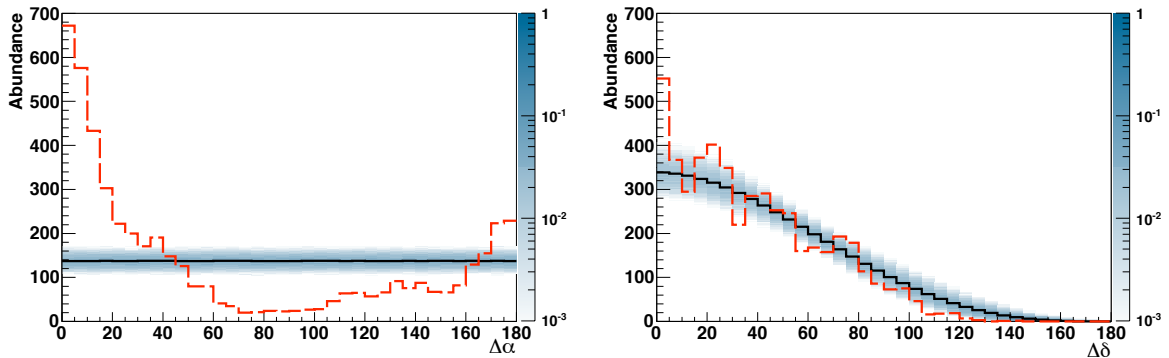


Figure 5.6: Reference Distributions of the TwoPtPlus method including the result for the test data set (red dashed). In black the mean expectation of isotropy and in blue shade the density distribution. *Left Panel:* Distribution of $\Delta\alpha$. *Right Panel:* Distribution of $\Delta\delta$.

5.2 Minimum Spanning Tree

In the 2pt-Correlation Function each pair of events A_i, A_j contributes to the estimator in the same way. All events are included independent of their relative distance towards each other, i.e. background events can dilute the anisotropy signal. This is especially the case when many sources contribute just a small amount to the total flux of UHECRs. In this way, no certain angular scale is enhanced leading to a deviation from isotropy (see the next chapter).

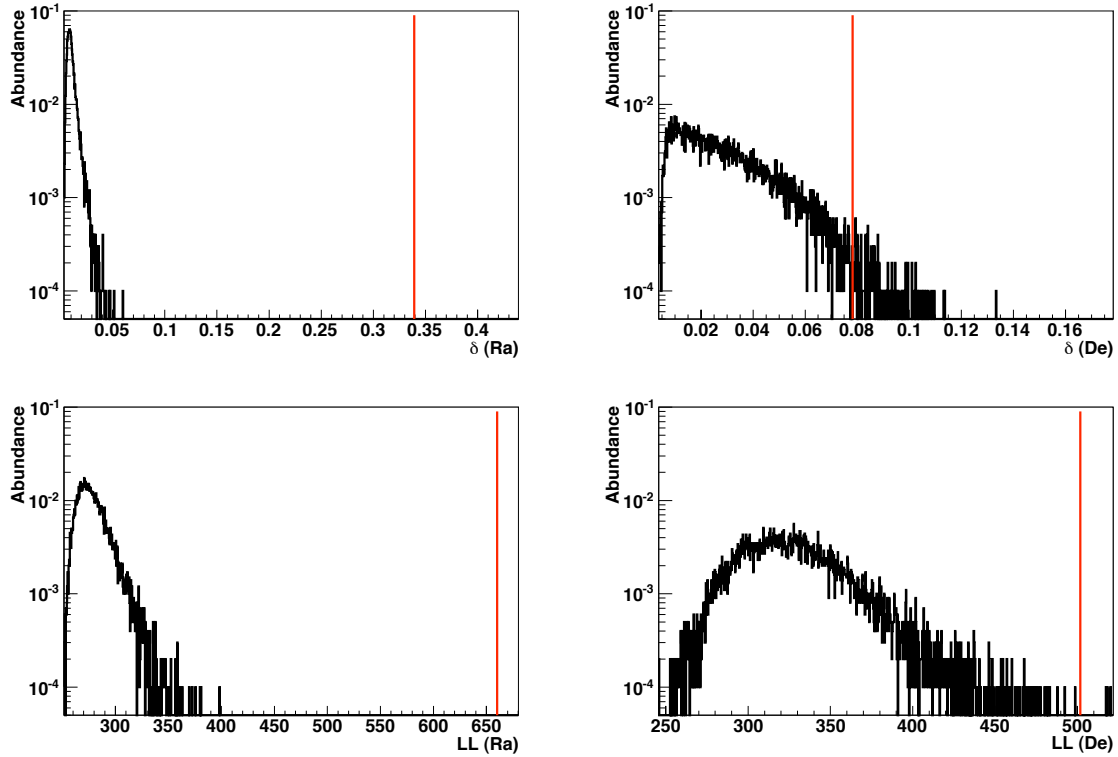


Figure 5.7: Estimator Distributions of the TwoPtPlus method including the result for the test data set (red line). In black the expectation based on 10,000 isotropy sets. *Top Left Panel:* Distribution of $\delta_{KS}(\Delta\alpha)$. *Top Right Panel:* Distribution of $\delta_{KS}\Delta\delta$. *Bottom Left Panel:* Distribution of $LL(\Delta\alpha)$. *Bottom Right Panel:* Distribution of $LL\Delta\delta$.

However, the influence of noise can be reduced by focusing only on adjacent events. A well known algorithm for such a purpose is the *Minimum Spanning Tree* (MST). The MST is an application from graph theory and has found its way into many areas of research amongst others astronomy [69] and astroparticle physics [70]. It connects a given set of n points along $(n - 1)$ weighted edges without loops minimizing the total cost or length of the tree, L_{MST} . In the case of distinct weights, the resulting tree is unique. Otherwise, several equal solutions are possible.

We adapt this concept for our autocorrelation studies and use the angular distance between two events as weight for the corresponding edge, L_e . The idea behind this is that for anisotropic source scenarios the events are distributed along the matter distribution forming large scale structures. In this case, many sources contribute a small amount of events each following the underlying source distribution which might be invisible in the estimator for the 2pt-correlation function. In contrast to this, the MST follows these kinds of structures resulting in a shorter length for the MST compared to the isotropic expectation.

To find the correct MST, we use Prim's algorithm which works as displayed in Fig. 5.8. Starting the MST with an arbitrary event, we search for another event connected to this along the shortest edge which is in turn added to the MST. Then,

the data set is searched for the next event close to one of the points already being part of the MST but this event must not be part of the MST. This process is repeated until all events are connected. A simple example consisting of four events is shown in Fig. 5.9. Note that the weight of an edge can differ from the geometrical distance between two events, i.e. additional factors can be included.

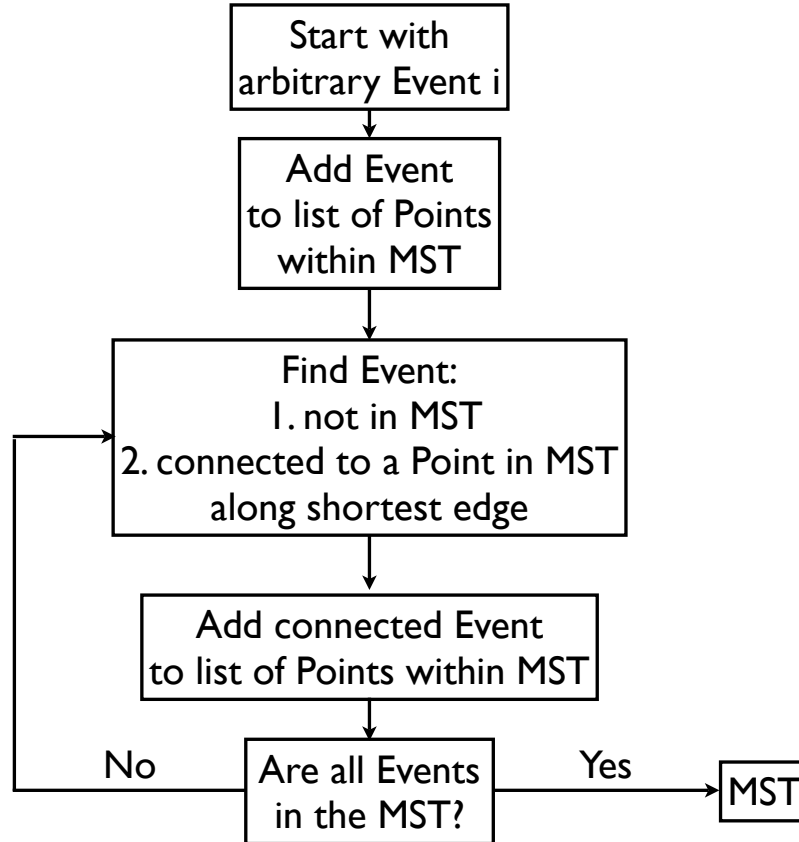


Figure 5.8: Starting a MST with an arbitrary Event i . Searching the data set for the next point which is not part of the MST and the closest to one point in the set. After adding this to the tree the algorithm is repeated until all events are part of the MST.

Applied to our example data set, the resulting MST on a sky map is shown in Fig. 5.10. Here, a filamentary structure is clearly visible with a nearly constant longitude of 180° . Such a behaviour could hint towards a possible source distribution when applied to the final data set. In Fig. 5.11 the results for the isotropic sets (black) and the test data set (red,dashed) are shown. On the left panel the edge length distribution, $L_{e,test}$, is shown affirming the clustering on small scales seen already by the 2pt-Correlation Functions. However, here it is much more enhanced since only the closest events contribute.

Again, the 10,000 isotropic sets provide the reference MST total length distribution (right panel). The analysis using the test data set yields a total length for the MST of $L_{MST,test} = 721$. By integrating from $L_{MST,test}$ to zero, a probability of $P < 10^{-4}$ is obtained that the events of the test data set are isotropically distributed.

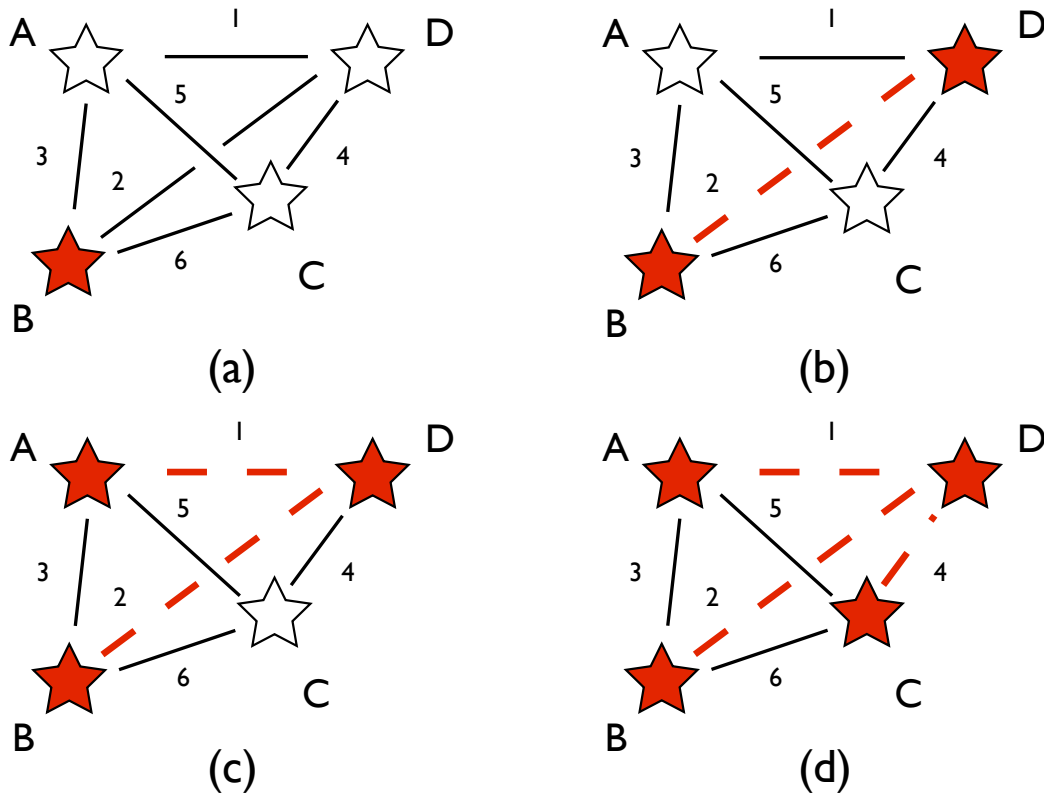


Figure 5.9: Example of the MST applied to a small event set. Points within the MST are marked as red stars and are connected via red dashed lines. Note that the geometrical distance between two events differs here from the actual weight of an edge. (a) Event B is the starting point of the Minimum Spanning Tree. (b) The edge with the smallest weight is found towards D. (c) Event A is closest to one point within MST, in this case again to D. (d) The last event is added again along the shortest edge starting at point D resulting in a total length for the MST of 7.

5.3 Cluster Algorithm

In contrast to the other methods, the Cluster Algorithm is designed to be sensitive to an overabundance of doublets (pairs of events with small angular distance), triplets and so forth in an event set, compared to an isotropic expectation. For this purpose, each event functions as a seed for the algorithm, i.e. starting with the first event (A_1) we calculate the angular distance Ω to all other events (A_j) in the set. This is done for all possible combinations resulting in a symmetric $n \times n$ -matrix O with the entries:

$$O_{ij} = \Omega_{ij} = \langle (A_i, A_j) \rangle. \quad (5.3)$$

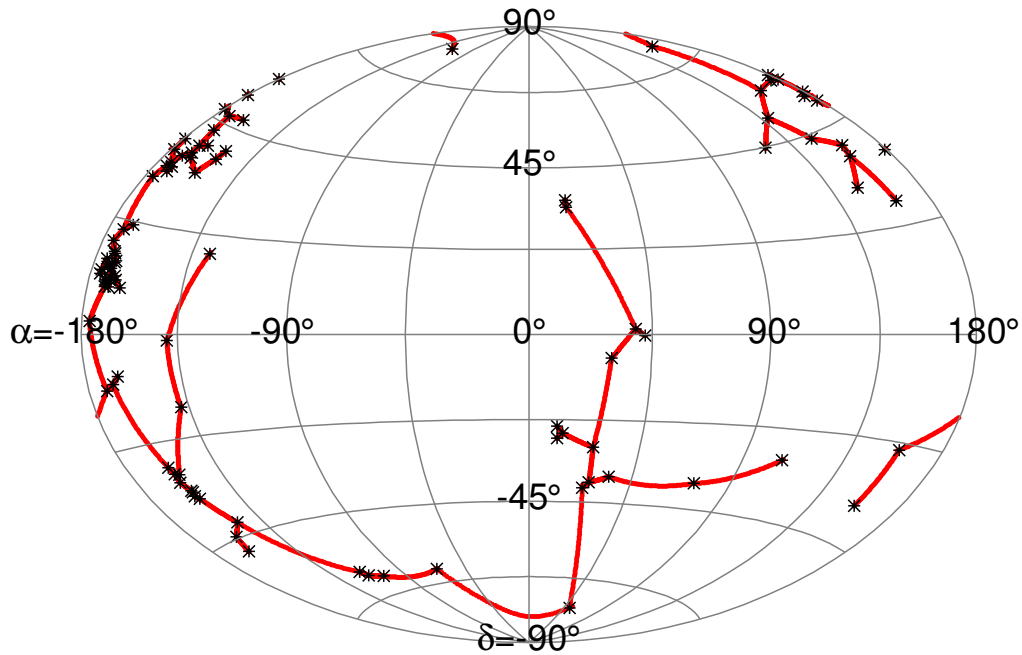


Figure 5.10: The example event set (black stars) with the corresponding Minimum Spanning Tree (red lines).

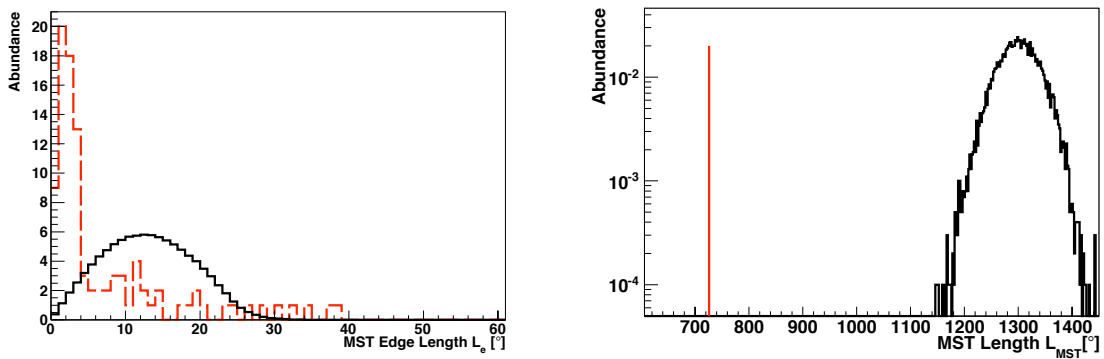


Figure 5.11: *Left Panel:* Distribution of edge lengths L_e within the MST for isotropy (black) and the test data set (red, dashed). *Right Panel:* Distribution of L_{MST} based on isotropy and the result for the data set (red line).

A simple example consisting of three events could lead to the following Matrix:

$$\text{e.g. } \begin{pmatrix} 0^\circ & 3^\circ & 4^\circ \\ 3^\circ & 0^\circ & 7^\circ \\ 4^\circ & 7^\circ & 0^\circ \end{pmatrix}.$$

However, at a certain angular scale the signal is diluted due to the increasing influence of the isotropic background. Therefore, we introduce a maximum angle R as a cutoff accepting only those entries where $O_{ij} \leq R$. The remaining multiplets (pairs, triplets, ...) are the found clusters.

To estimate a reasonable value for R we use a formula given in [53] based on the random walk model (see chapter 3.3.3.2):

$$R \simeq 0.025^\circ Z \left(\frac{d}{\lambda} \right)^{1/2} \left(\frac{\lambda}{10 \text{Mpc}} \right) \left(\frac{B}{10^{-11} \text{G}} \right) \left(\frac{E}{10^{20} \text{eV}} \right)^{-1}. \quad (5.4)$$

Assuming protons as primary particles ($Z=1$) with an energy of $E = 55 \text{ EeV}$, $\lambda = 1 \text{ Mpc}$ the typical size of a region with a coherent magnetic field $B = 2 \text{ nG}$ and $d = 100 \text{ Mpc}$ the maximal allowed distance between source and detector leads to an estimated maximum angular distance of $R \approx 10^\circ$.

In the following, we present two variants of the algorithm. While the first one is focusing solely on the number of events within a cluster, the second one includes also information about the distribution of events within a cluster itself. In both cases, we demonstrate the performance of the algorithm by means of the example event set (see Fig.5.12). Of the found clusters we only display doublets and triplets by red circles around the centre of each cluster with radius R due to reasons of better clarity.

It is obvious that clusters overlap. The reason for this is, that if two events A_i and A_j are closer than R then accordingly this statement is true vice versa, i.e. each doublet is found twice. However, for multiplets of higher order this is not always true, since on average the events are further apart so that a triplet includes more often two doublets than two additional triplets. In the end, all clusters are kept leading to the conclusion that the result of the algorithm is independent of the starting event.

5.3.1 Unweighted Realisation (CA I)

Without making any assumptions about distributions of UHECR around their sources, we just account for the events within a cluster in the same way by setting all elements O_{ij} surviving the maximum angular distance cut to one and the rest to zero. This includes also the entries on the diagonal ($i=j$). This new matrix is called M . Two simple examples are shown in Fig. 5.13 consisting of three events each. In the first example (a), only the events 1 and 2 are closer to than R to each other while the angular distances of event 3 to the others is larger than R . In the bottom panel (b), all events are part of a cluster where event 1 and 3 just close to event 2.

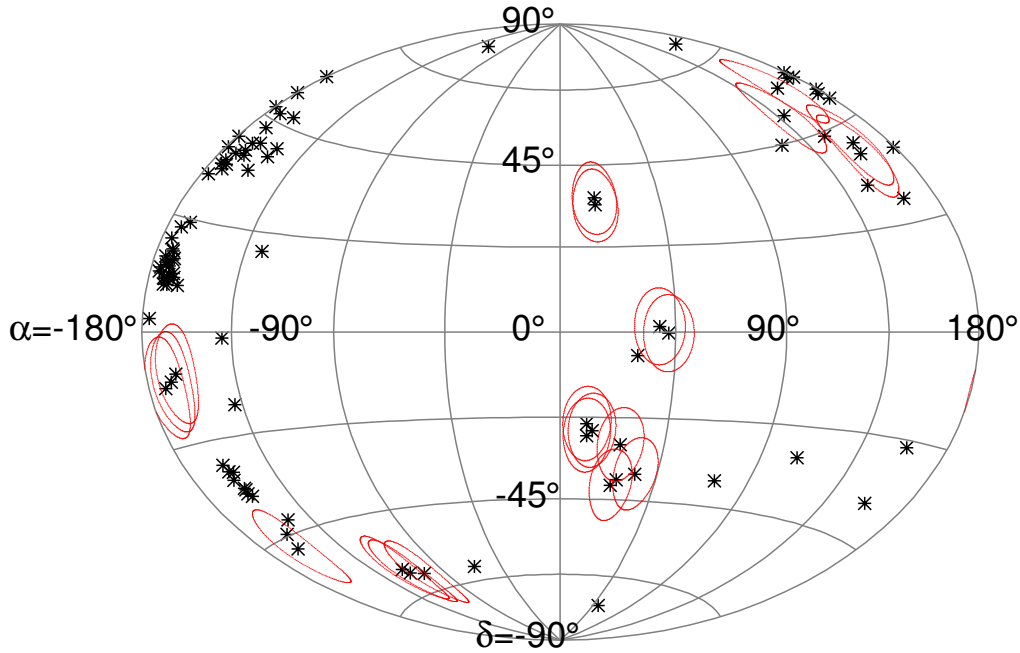


Figure 5.12: An example Monte Carlo data set (100 events) as black stars. Only the found doublets and triplets are marked as red circles drawn around the seed event due better clarity.

Considering these examples, the first step regarding the calculation of the estimator is to sum over all entries in row i of matrix M resulting in the Cluster Size, $N_{C,i}$, which corresponds to the type of found cluster (pair, triplet, ...). This sheds light on the abundance of isolated events, doublets and so forth. The final estimator is the Cluster Sum N_C , the sum over all entries of the matrix M . Following these rules, example (a) gives $N_C = 5$ and example (b) $N_C = 7$.

In the following the CA I is applied to the example event data set. The left panel of Fig. 5.14 contains the distribution of the Cluster Size for the test data set (red, dashed) and the isotropic expectation (black). While for isotropy the isolated events and doublets dominate as expected, the test data set differs clearly from H_0 .

$$N_C = \sum_{i=1}^n \sum_{j=1}^n M_{ij}(R) = \sum_{i=1}^n N_{C,i}. \quad (5.5)$$

The larger N_C gets the more anisotropic is the analysed data set, depending on the choice of R . From this it follows that N_C is much larger for the test data set as for the isotropic expectation based on 10,000 sets as can be seen in the right panel of Fig. 5.14. Though, if R is too large, for instance close to 180° , then N_C will be similar for all scenarios.

Finally the method yields a value of $N_{C,test} = 864$ for the test data set representing a probability of $P < 10^{-4}$ to be in agreement with the isotropic expectation.

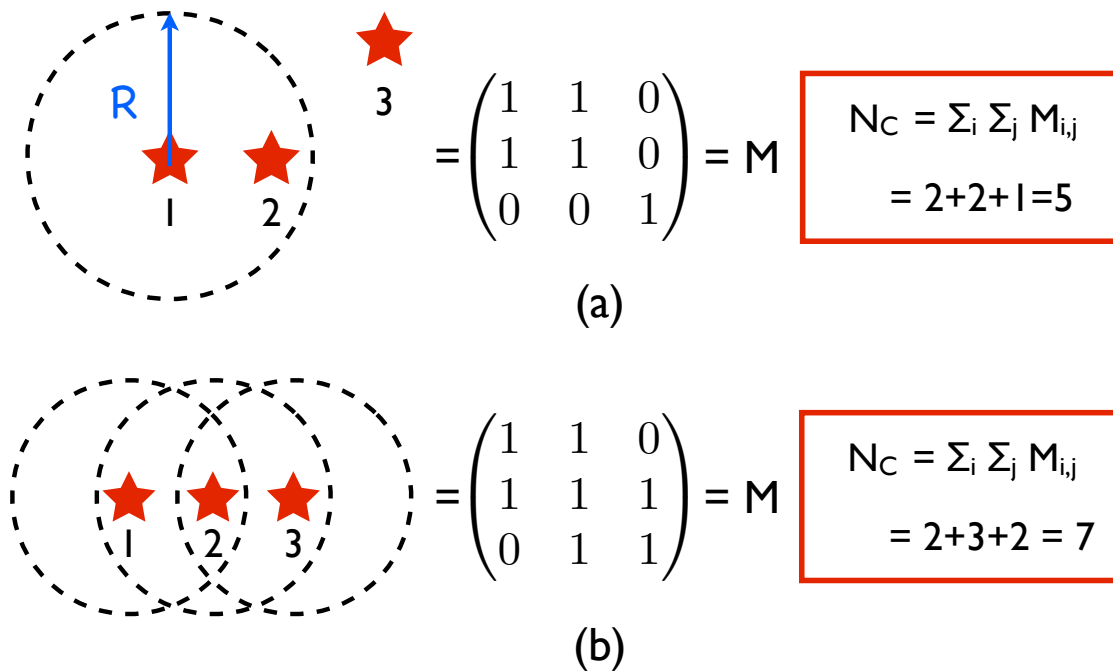


Figure 5.13: Two simple applications of the CA I to small events sets are shown with three events (red stars) each. (a) While the events 1 and 2 are closer to each other than R , event 3 is too far away. Thus it contributes to no cluster. The Cluster Size $N_{C,i}$ and the Cluster Sum N_C are calculated by summing over row i of matrix M and by summing over all entries of matrix M , respectively. (b) Event 2 is closer than R to the other events forming a triplet while the others in turn only form a pair with event two each. The calculation of the estimator is the same as before.

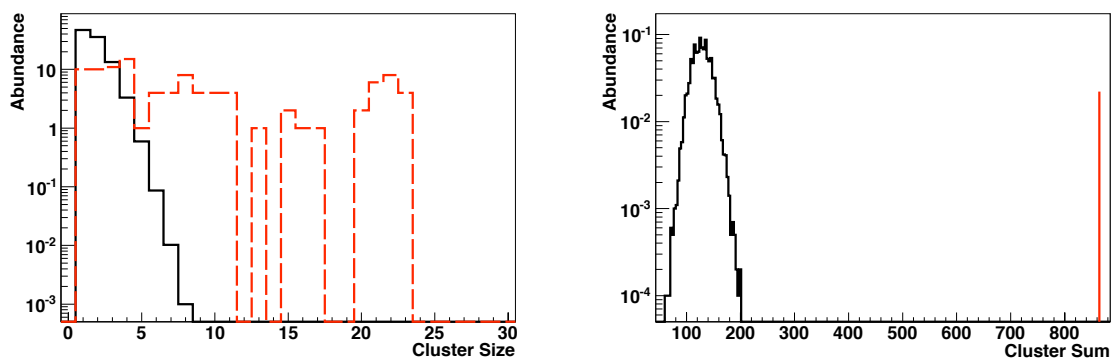


Figure 5.14: *Left Panel:* Distribution of Cluster Size for 100 events: Isotropic expectation in solid black for 10,000 sets normalised to one set and test data in dashed red. *Right Panel:* Distribution of Cluster Sum for 100 events: isotropic expectation in black for 10,000 sets normalised to one and test data set (red).

5.3.2 Weighted Realisation (CA II)

In contrast to the unweighted variant, a weight w can be attributed depending on the angular distances Ω of the events within clusters, which should be large for small Ω 's and close to zero for $\Omega \approx R$, since for anisotropic scenarios it is expected that events are closer to each other than in the isotropic case. Here, we adapted a Gaussian distribution to calculate the weight motivated by the random walk model [53] for UHECR predicting this behaviour.

$$w_{ij} = \exp\left(-\frac{\Omega_{ij}^2}{2\sigma^2}\right) \quad (5.6)$$

$$\sigma = \frac{1}{3}R \quad (5.7)$$

Thus, the above mentioned requirement is fulfilled having a weight close to one for nearby events and one close to zero if their distance is comparable to R . Furthermore, the standard deviation σ is directly linked to the previously defined maximum radius keeping the number of free parameters fixed to one.

The algorithm uses the also the original matrix O containing all relative angular distances between all events which are smaller than R . However, instead of setting all entries unequal zero to one as for CA I, here the entries smaller than R are set to $w_{i,j}$ and the rest is zero. In addition, the entries on the diagonal are all also set to zero since they only add a constant offset to N_C which is equal to the size of the event set. Again, the resulting matrix is called M . In equation (5.5) the Cluster Size $N_{C,i}$ is replaced by the Cluster Weight $W_{C,i}$ which is the sum of all entries of row i of M . Then we calculate the Cluster Weight and Cluster Sum as defined in equation (5.5). Besides the information about the abundance of each Cluster Size, Fig. 5.15 also includes colour-coded the distribution of the weights for all occurring Clusters Sizes. Note that the Cluster Size is still an integer as defined in the previous section while the Cluster Weight is a floating number.

As before, the isotropic model gives mostly values $N_{C,i} = 2$ with small weights, i.e. their angular distance $\Omega_{ij} \approx R$ (use equation (5.6)) which is a new information inaccessible in CA I. Again, the difference between model and test data set is obviously leading to a similar plot as shown for the unweighted case (see right panel of Fig. 5.14). The Cluster Sum (CA II) $N_{C,test}$ is 295 for the example data set with the corresponding probability of $P < 10^{-4}$ (see Fig. 5.16). In the case of CA II the Cluster Sum is smaller than for the unweighted case, which is expected. Since the Cluster Weights are in general smaller than 1, while the Cluster Size is at least 1.

5.4 Summary

In this chapter three different algorithms have been presented. Each one has different assets and drawbacks by design. Here, we have explained in detail how these work by means of a very anisotropic test data sample. The single steps have been presented resulting in the corresponding estimator. All methods have found a clear deviation

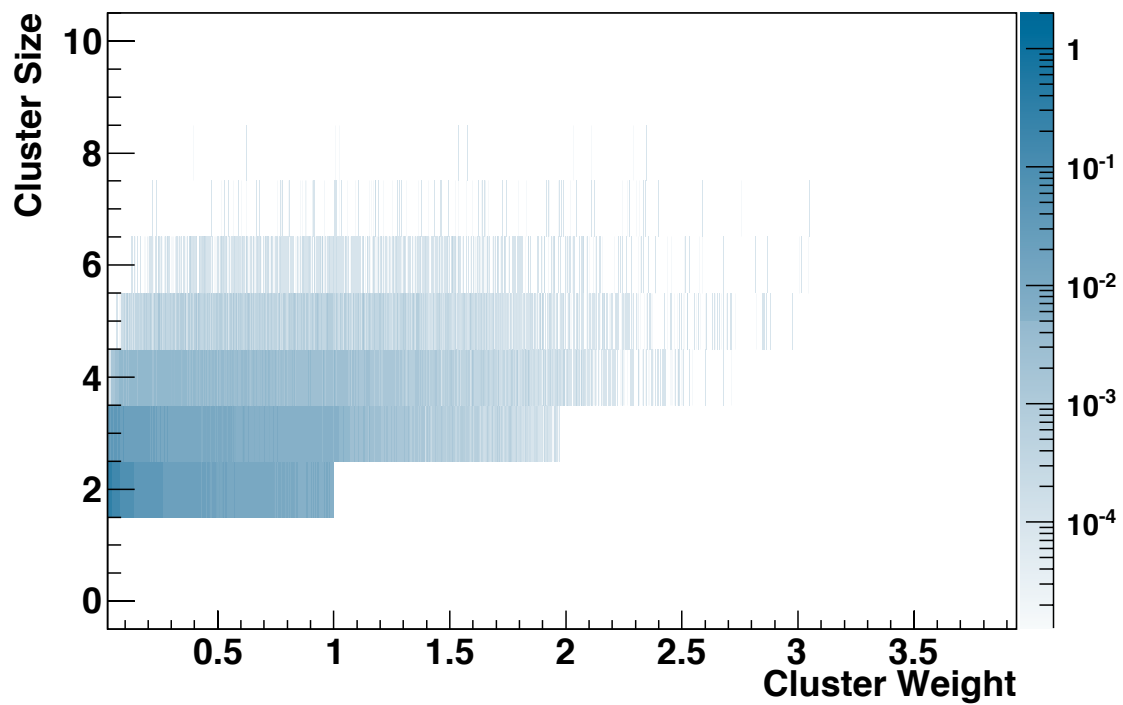


Figure 5.15: Distribution of Cluster Size for 10,000 isotropic sets of 100 events. Doublets with small weights dominate.

from the Null-hypothesis, an isotropic distribution of arrival directions of UHECRs. A more detailed analysis regarding the sensitivity is performed in the next chapter by using several different source scenarios and testing the methods on the resulting Monte Carlo sets.

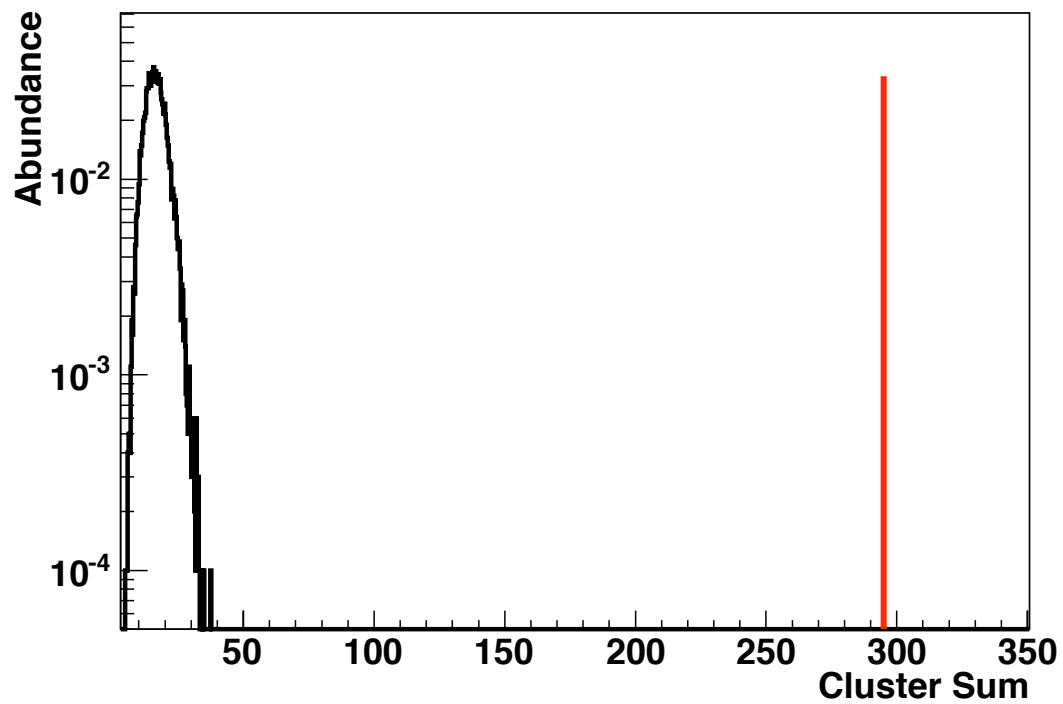


Figure 5.16: Displayed is the distribution of the Cluster Sum for 10,000 isotropic sets containing 100 events each. The test data set (red solid line) shows a clear deviation from the isotropic expectation yielding a probability of 10^{-4} .

6. Monte Carlo Studies

Before applying the autocorrelation methods to the data set of the PAO, we explore their ability to find a deviation from isotropy. Therefore, many Monte Carlo data sets were simulated varying the underlying type and amount of anisotropy as well as the total number of events. An advantage of this procedure is the possibility to recognize cross-correlations among the methods and study the effect of the PAO exposure in contrast to a full sky observatory. We start by discussing the procedure of generating these Monte Carlo Maps commonly known as *Mock Maps* of the considered anisotropies and end with the results obtained by analysing them with the methods presented in the previous chapter.

It is worth noting that in the following no detector simulations are included and thus, the anisotropic MC sets are assumed to be ideal. For isotropy, the detector effects are negligible due to the fact that they would lead to a further isotropization of the sample. However, the reference is already totally isotropic and therefore, it would be unaltered.

6.1 Generating Mock Maps

Despite the vast amount of possibilities regarding source scenarios of UHECRs, we focus on three basic types here:

1. catalogue-based (many sources),
2. isolated objects (few sources) and
3. multipoles.

Although this choice covers current and most prominent source models, it is impossible to consider all potential scenarios, e.g. strengths and number of sources as well as their distribution and the scale of deflection.

In principle, the analysis procedure is the same for all source scenarios. The unit sphere which is the reference for the arrival direction has to be binned and each bin has the weight proportional to the probability $\rho = \rho(\phi, \theta)$ that cosmic rays originate from that position in the sky. This depends of course also on the type of smearing and the smearing scale applied.

Then, 10,000 sets each containing from 20 up to 100 cosmic ray events are produced according to the density distribution of these maps using a step size of 20. This choice of the number of events reflects the current statistics available from different

experiments, e.g. the PAO has above 55 EeV roughly 90 events up to now. The resulting density maps shown below are all in equatorial coordinates displayed according to the Hammer projection scheme.

It is worth mentioning, that standard programs like *ROOT* cannot bin the sphere in an optimal way. The typical rectangular binning of a 2-dim histogram matches different solid angles to the same area. One possibility to avoid this problem is the usage of the HEALPix package (see section 6.1.1) provided by the Cosmic Microwave Background community.

6.1.1 HEALPix

The name HEALPix stands for High Equal Area isoLatitude Pixelisation [71] and solves the issue of different solid angles per pixel. This is achieved by using curvilinear quadrilaterals (cf. Fig. 6.1) covering equal area partitions of the sphere. Each one of them can in turn be split into four smaller quadrilaterals having the same properties as their mother pixels. In the basic configuration the corresponding scheme consists of twelve pixels ($= N_{pix}$) with a resolution of three pixels in θ and of four pixels in ϕ . Instead of the the number of pixels, the describing parameter is N_{side} which is here equal to 1. With each quartering of the pixels N_{side} is increased by a factor of two such that the following formula per definition describes the total amount of pixels on the sphere:

$$N_{pix} = 12 \cdot N_{side}^2. \quad (6.1)$$

For $N_{side} = 1, 2, 4, 8$ the realisations are shown in Fig. 6.1 starting at the upper left and then going clockwise to higher N_{side} . The black dots represent the center of the pixels and as one can see, they are arranged in rings of constant latitudes.

6.1.2 Smearing

To create realistic density maps for the different source scenarios, it is necessary to include deflections caused for example by intergalactic magnetic fields. The simplest model is called random walk [44] assuming many regions of non-coherent fields between origin and detector. This leads to a Gauss distribution of events around their source.

The effect of the directional galactic magnetic field B_G is neglected here. Since, one can argue that due to the high energies of the considered cosmic rays and the small amount of time, UHECRs are only minimally affected by B_G . Thus, the deflection is small compared to the one caused by the intergalactic magnetic field.

On a 2-dimensional sphere, the Gaussian is replaced by the von Mises-Fisher distribution for the \mathfrak{R}^3 :

$$f_3(\mathbf{x}; \mu, \kappa) = \frac{\kappa}{2(e^\kappa - e^{-\kappa})} \cdot \exp(\kappa \mu^T \mathbf{x}). \quad (6.2)$$

Here, μ is the position direction vector of the source and \mathbf{x} the vector to the point on the sphere where one is interested in the probability density. Thus, the exponent

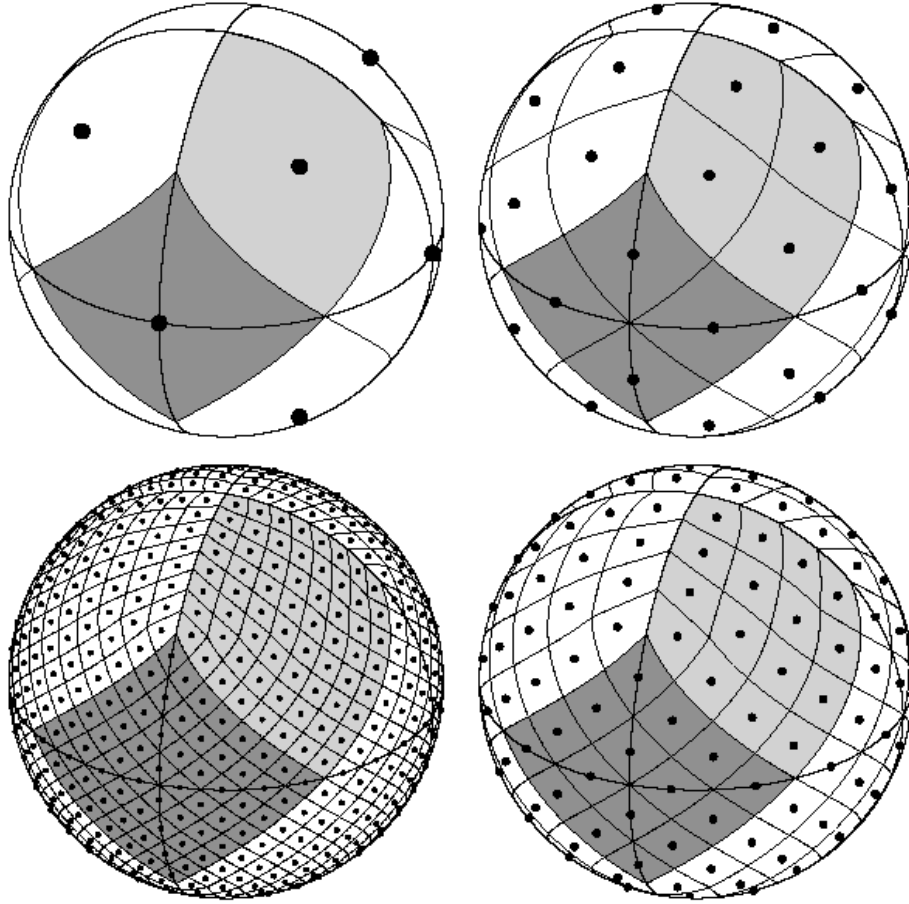


Figure 6.1: Displayed is the pixelisation of the sphere for $N_{side} = 1, 2, 4, 8$ starting at the upper left panel and then going clockwise (taken from [71]). The sizes which are given in the following represent the average radius of a quadrilateral. *Upper Left Panel:* $N_{side} = 1$ leads to 12 pixels, each a size of $\approx 45^\circ$. *Upper Right Panel:* $N_{side} = 2$ leads to 48 pixels, each a size of $\approx 22^\circ$. *Bottom Right Panel:* $N_{side} = 4$ leads to 192 pixels, each a size of $\approx 11^\circ$. *Bottom Left Panel:* $N_{side} = 8$ leads to 768 pixels, each a size of $\approx 5^\circ$.

can be simplified. It is just the vector product of these two vectors resulting in the cosine of the angle Ω in between.

$$\kappa\mu^T \mathbf{x} = \kappa \cdot \cos \Omega. \quad (6.3)$$

For small deflections the von Mises-Fisher distribution can be approximated by a Gaussian leading to the relation $\kappa = 1/\sigma^2$.

The binning is here important, too. Since, the differential angular distance should be unaltered to ensure a consistent behaviour regarding the random walk model, i.e. the angular distance distribution of events around the source should follow a Rayleigh distribution.

Therefore, an N_{side} of 256 is chosen with a pixel radius of $\approx 0.2^\circ$ well below the 1° angular resolution of the PAO. Thus, artificial deviations from the expected Rayleigh distribution caused by the binning around a source vanish.

6.1.3 Effect of Coverage

Until now, the complete sphere has been considered. However, in this analysis the data of the Pierre Auger Observatory is only considered up to a zenith angle θ smaller than $\theta_m = 60^\circ$ in local coordinates leading to a sharp cut at a declination of $\delta \approx 25^\circ$. The resulting coverage ω depending on the declination δ is analytically described in [72] :

$$\omega \propto \cos(\delta) \cos(\delta_A) \sin(\alpha_m) + \alpha_m \sin(\delta) \sin(\delta_A) \quad (6.4)$$

where $\delta_A = -35.2^\circ$ is the declination of the PAO and α_m is

$$\alpha_m = \begin{cases} 0 & \text{if } \xi > 1 \\ \pi & \text{if } \xi < -1 \\ \cos^{-1} \xi & \text{otherwise} \end{cases}$$

with

$$\xi \equiv \frac{\cos \theta_m - \sin \delta_A \sin \delta}{\cos \delta_A \cos \delta}.$$

It must be noted that this formula is independent of the right ascension.

This is a parametrisation of isotropy measured by the PAO and a density HEALPix map can be filled according to this formula (see Fig. 6.2). By multiplying this reference with an anisotropic density map we obtain the probability density distribution for the arrival directions of cosmic rays as seen by the PAO for this scenario.

6.1.4 Catalogues

Many catalogues exist in astronomy mostly focusing on a single wavelength band, e.g. radio, x-ray and infrared, or on special objects, for example Active Galactic Nuclei (AGN). The second type of catalogue was used by the PAO for the catalogue-dependent anisotropy study described in [26]. Therefore, we use the same one, the Veron-Cetty and Veron (VCV) catalogue to generate MC maps. However, it has some known issues which are discussed below. As a contrast to the VCV, we have chosen the Infrared Astronomical Satellite (IRAS) catalogue in addition which fulfils the requirement of being complete up to a distance of 100 Mpc (see [73]).

6.1.4.1 VCV

The Veron-Cetty and Veron (VCV) catalogue [3] is a compendium of many different objects. It contains over 10,000 objects consisting of AGN as well as BL Lac and quasar candidates which can be found in the literature. The coverage of the catalogue is non-uniform which can be seen in the top left panel of Fig. 6.3. Obviously a band is missing with longitudes of roughly $\pm 90^\circ$. The reason for this lies in the presence of the milky way obscuring the sight on those AGN lying behind it.

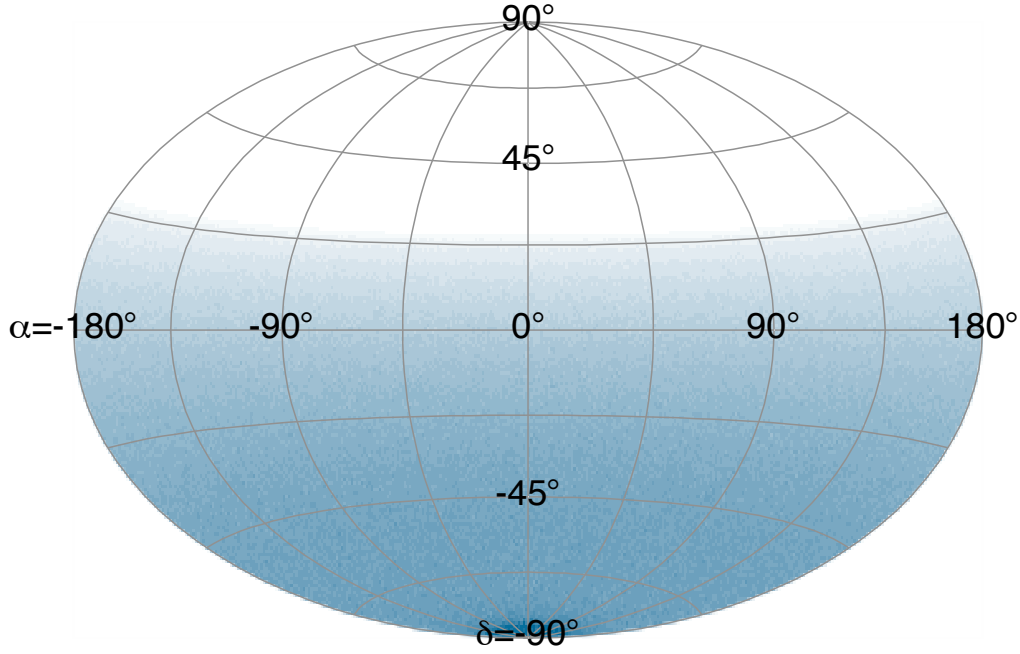


Figure 6.2: The analytically calculated coverage of the Pierre Auger Observatory. The highest probability for events to be detected is at the south pole since it is in the field of view of the observatory at all times.

However, on this very catalogue the correlation analysis of the PAO was performed with a restriction on the redshift of the objects ($z \leq 0.017$ which corresponds to 428 objects on the whole sky).

Therefore, we use this catalogue as a basis for one source model whereupon we extend the allowed range of redshift to 0.02 (≈ 85 Mpc including 497 objects, cf. top right panel of Fig. 6.3). Since the redshift cut obtained during the correlation analysis is the result of a scan for this catalogue-dependent approach and was optimized within the framework of that study. Thus, the sources still might have larger distances than 74 Mpc. The GZK-effect for instance predicts that cosmic rays at the highest energies have their origin within distances of 100 to 200 Mpc. As a compromise between both values we have chosen the one mentioned above.

In addition, we have chosen two kinds of strength assignments for each source. The first one assumes that all objects contribute to the same amount of cosmic rays on Earth independent of their distance to us. This corresponds to different strengths of each object. The second one assigns to each source candidate a weight which is proportional to $1/z^2$, i.e. each object has the same absolute magnitude. For close objects the linear Hubble law can be applied where $z \propto d$.

For the smearing scale we adopt another result of the correlation analysis namely the maximal angular distance being $\approx 3^\circ$. Furthermore, we increase this number in another event set to 6° to estimate how the signal might behave when the deflections is larger.

After the amount and relative strength of the sources including the smearing scale is defined, we set in the last step the signal to noise ratio per event set. We vary the isotropic contribution from 0 % to 80 % in steps of 20 %.

In the bottom panel of Fig. 6.3 an example is shown based on a pure VCV contribution weighted with the inverse of the square of the distance. On the left side of the bottom panel the pure distance weighted VCV catalogue with a smearing of 3° is displayed which gives, by multiplying it with the PAO coverage, the map on right. The effect of the coverage is clearly visible, particularly close to the south pole where the sources are enhanced in comparison to those in the vicinity of the equator. Further maps based on different parameters can be found in the appendix.

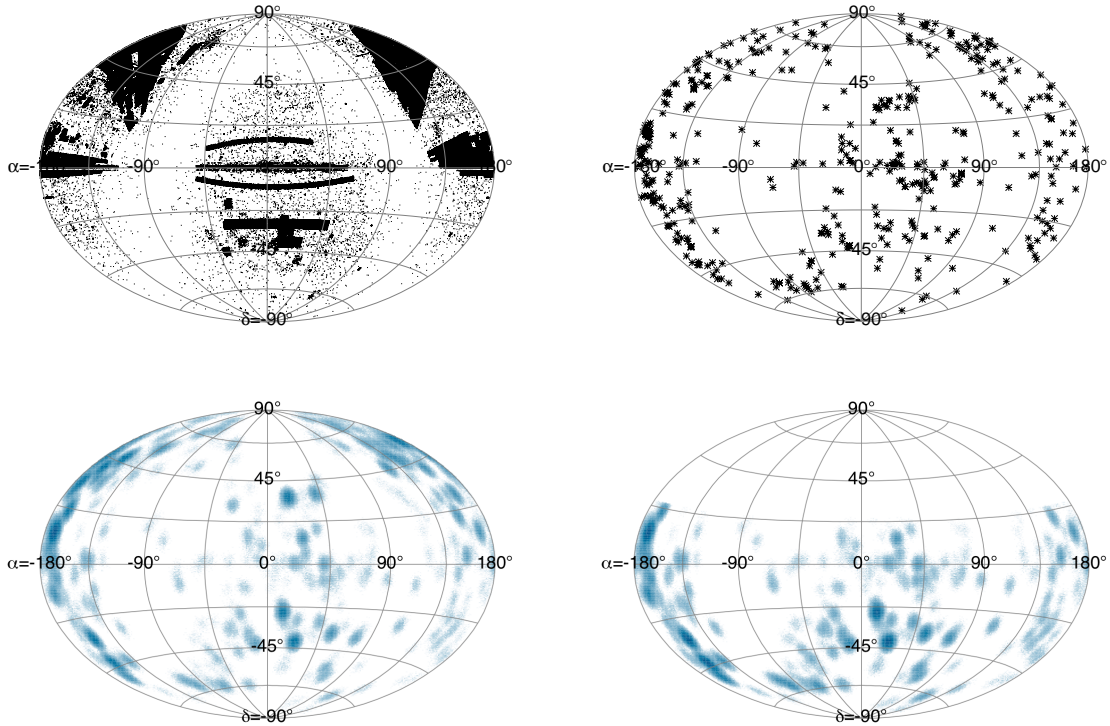


Figure 6.3: The Veron-Cetty and Veron catalogue in different representations in equatorial coordinates and Hammer projection. The strength of the source is indicated coloured plots, the darker the blue the stronger the source. The logarithmised strength is given in arbitrary units. *Top Left Panel:* Displayed are all objects having a redshift. *Top Right Panel:* The 497 objects are shown having a redshift smaller than 0.02. *Bottom Left Panel:* These 497 objects are smeared with a Gaussian ($\sigma = 3^\circ$) and weighted with the inverse of the distance squared. *Bottom Right Panel:* The density map on the left side for the complete sphere is weighted with the coverage of the PAO enhancing objects close to the south pole while those close the border of the field of view are suppressed.

6.1.4.2 IRAS

The Infrared Astronomical Satellite (IRAS) [74] was an experiment operating for ten month in 1983 searching the sky for infrared sources. It detected the positions

of more than 200,000 objects, and the distances are not measured.

In the following years dedicated redshift surveys determined the corresponding distances of several thousand of these objects. A combination was compiled by [75] taking care of uniform sky coverage and completeness of the final catalogue, the Point Source Catalogue redshift (PSCz).

However, some regions of the sky are more trusted than others due to unknown exposure in that region of the sky. These unreliable regions are masked to ensure a reliable distribution of sources. The mask applied here is called *himask* and yields finally than 10,000 objects (see top left panel of Fig. 6.4). Comparing this map with the one of the VCV a clear difference emerges. Indeed, the PSCz seems to be much more uniform although the same empty band is present in the sky as before again caused by the milky way obscuring objects lying behind.

To generate Mock Maps based on this catalogue, the same specifications are used as in the case of the VCV, i.e. $z \leq 0.02$, smearing 3° and 6° , respectively, and a flat distribution as well as a distance weighting for the strength of the sources. In addition, the same contribution of an isotropic background is chosen.

The example map ($\sigma = 3^\circ$, distance weighting) is shown on the bottom left panel which is dominated by two sources in the northern hemisphere close to the pole. By multiplying it with the coverage of the PAO the final density map is obtained whereupon now many sources are becoming visible on the southern hemisphere.

It is worth noticing that the general matter distribution of the VCV and PSCz is following the same directions.

6.1.5 Single Sources

A complementary approach focuses on single sources. As described in chapter 3.3.1.1, a model based on radio loud galaxies proposes just this kind of anisotropy. Few sources contribute a major part to the overall flux of UHECRs which are more or less deflected by magnetic fields.

In addition, it is still possible that the current source scenarios are all incorrect, i.e. in general without making any assumption all positions are equal. Therefore, we introduce here a scenario where the sources are randomly distributed on the sky.

In both cases the signal is diluted by a varying amount of isotropic background (40 %, 60 %, 80 %).

6.1.5.1 Radio Loud Galaxies

The three radio loud galaxies mentioned before are considered for this model, namely Centaurus A, Virgo A, Fornax A. Compared to the catalogue-based scenarios, a more explicit model of strength and smearing is possible. We use two kinds of smearing

1. Gaussian (based on random walk [44]) and
2. and Inverse (proposed by P. L. Biermann, see below).

The next step is to fix the relative strength of each source. We use two models suggesting different mechanisms powering the jets which in turn accelerate the particles

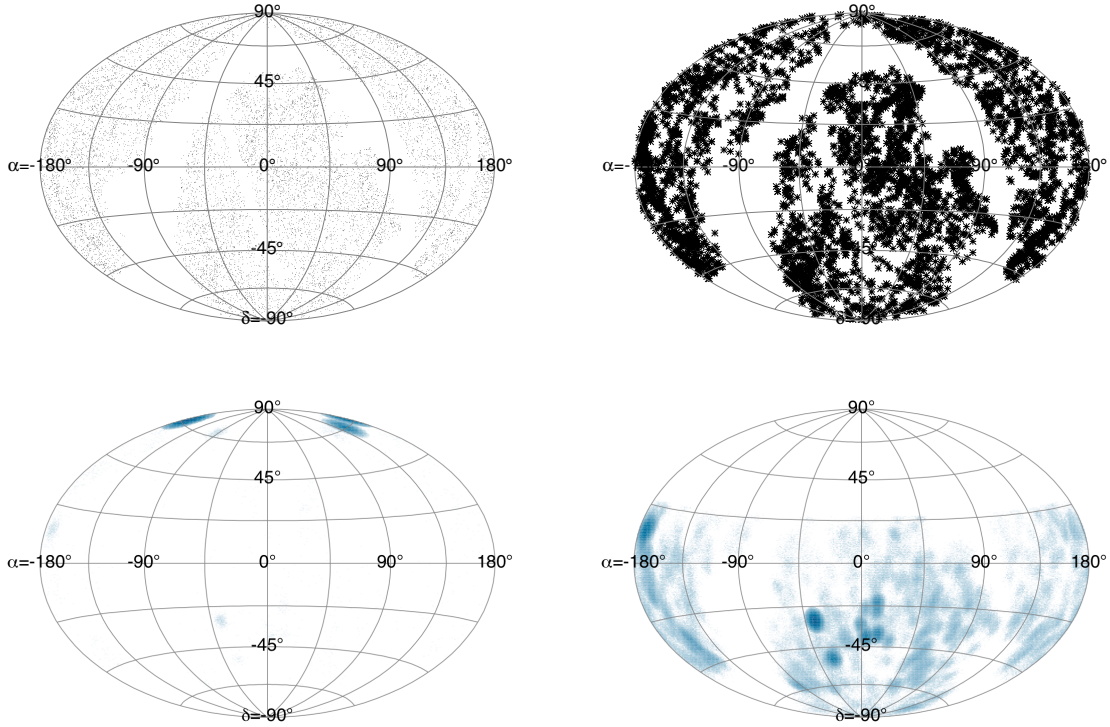


Figure 6.4: The Infrared Astronomical Survey catalogue in different representations in equatorial coordinates and Hammer projection. Regarding the coloured plots, the darker the blue the stronger the source. The logarithmised strength is given in arbitrary units normalised to the same value. *Top Left Panel:* Displayed are all objects within the catalogue. *Top Right Panel:* The 4431 objects are shown having a redshift smaller than 0.02. *Bottom Left Panel:* The remaining objects are smeared with a Gaussian ($\sigma = 3^\circ$) and weighted with the inverse of the distance squared. Three sources on the northern hemisphere dominate the total flux. *Bottom Right Panel:* The density map on the left side for the complete sphere is weighted with the coverage of the PAO enhancing objects close to the south pole while those close to the border of the field of view are suppressed.

within. The first one is the Spin-Down model introduced by Blandford and Znajek in 1977 (model S, [76]) and the second is the Accretion model by Falcke (model A, [77]). Both allow to calculate the expected relative flux on earth for each galaxy, which has been computed by P. Biermann [29] (see Tab.6.1).

Gaussian Model

We calculated the smearing width σ according to formula 6.5 which is valid for this kind of model:

$$\sigma = 0.025^\circ \left(\frac{d}{\lambda} \right)^{1/2} \left(\frac{\lambda}{10 \text{Mpc}} \right) \left(\frac{B}{10^{-11} \text{G}} \right) \left(\frac{10^{20} \text{eV}}{E} \right) Z. \quad (6.5)$$

d is the distance of each object to the Earth, $\lambda=100$ kpc the average coherence length of a magnetic field with strength B and E the energy of the cosmic ray with

Table 6.1: Predictions of relative fluxes for Models S and A calculated by P. Biermann in [29].

Name	relative strength	
	Model S	Model A
Cen A	13	3.63
Vir A	1	1
For A	0.01	0.03

charge Z . We adopted the lower energy bound used for the AGN correlation study, i.e. $E=55$ EeV and set $B = 10^{-9}$ G, which is within the upper limit on the strength of the intergalactic magnetic field (1 - 10 nG) set by recent CMB measurements [42]. By increasing the average coherence length to 1 Mpc or the strength of the B-field to 10 nG, σ is enlarged by a factor of ten. This leads to huge smearing scales and the arrival directions are isotropized. Therefore, we have chose the above mentioned values to generate a model which is able to produce the observed overabundance close the Cen A. The corresponding scale is of the order $10^\circ - 20^\circ$ (cf. Fig. 3.7).

Using the distances given in Tab. 3.2 the following approximate σ 's are obtained:

$$\sigma_{CenA} = 11^\circ \quad \sigma_{VirA} = 23^\circ \quad \sigma_{ForA} = 25^\circ \quad (6.6)$$

$$(6.7)$$

The maps based on acceleration model A with an isotropic background of 40 % are displayed in the top panel of Fig. 6.5 whereupon on the left the full sky map can be seen and on the right the one weighted with the coverage of the PAO. As one can conclude from the source strength given in Tab. 6.1, Cen A is the strongest clearly visible, followed by Vir A. Whereas, For A is invisible being on the same level as the isotropic background.

Inverse Model

The basic assumption in this model is that the smearing follows a $1/\Omega$ -distribution caused by a magnetic plasma. UHECRs mainly scatter just once in this environment and stay afterwards unaffected during propagation. In this way, the UHECRs are spread evenly over the complete sky with a maximum around the source (see [29]). This is an alternative approach which stands in contrast to the default distribution (Gaussian) based on the random walk model where many scatterings are taking place during the propagation of cosmic rays. Ω is again the solid angle between the source and a point on the sphere. The integral of the function, however, is unequal to one. Hence, in principle it cannot be used to represent a probability density. Therefore, the function has to be modified to ensure a proper behaviour. One option is to introduce a cut-off parameter $\epsilon > 0$ in the sense that

$$f(\Omega; \epsilon) = \frac{1}{N} \frac{1}{\Omega + \epsilon},$$

with N being the normalisation constant. We chose ϵ according to the σ 's computed in the previous section. The idea is to keep the same amount of events within this

predefined radius around the source. To determine the value of ϵ , we integrate $f(\Omega)$ only up to the corresponding smearing scale σ calculated for the Gaussian model which should yield a value of 68 %. By solving the following equation numerically, the ϵ is calculated.

$$\sigma = \left(\left(\frac{180}{\epsilon} - 1 \right)^{0.68} - 1 \right) \cdot \epsilon.$$

The different functions according to the corresponding strength and cut-off parameters are displayed in Fig. 6.6.

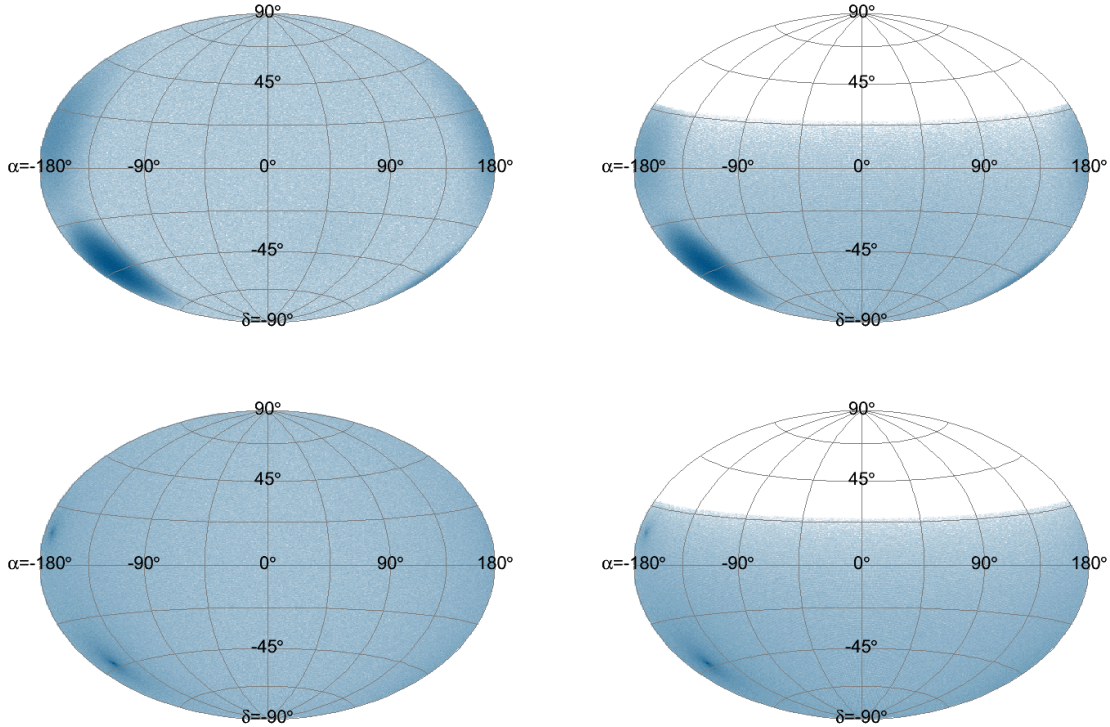


Figure 6.5: Probability density maps based on three radio loud galaxies and strength model A, the isotropic background is 40 %. The logarithmised strength is given in arbitrary units. *Top Left Panel:* The distribution on the complete sphere with an applied Gaussian smearing. *Top Right Panel:* Resulting map multiplied with the exposure of the PAO. *Bottom Left Panel:* The distribution on the complete sphere with an applied smearing according to the Inverse model. *Bottom Right Panel:* Resulting map multiplied with the exposure of the PAO.

The assigned probability for each pixel is the sum based on the values of the three curves corresponding to the three sources according to their distances to the sources. In the bottom of panel Fig. 6.5 we show the probability density of model A again with 40 % background. The sources are nearly invisible, only Cen A can be identified. However, the area which is enhanced is much smaller than the one seen for the Gaussian smearing.

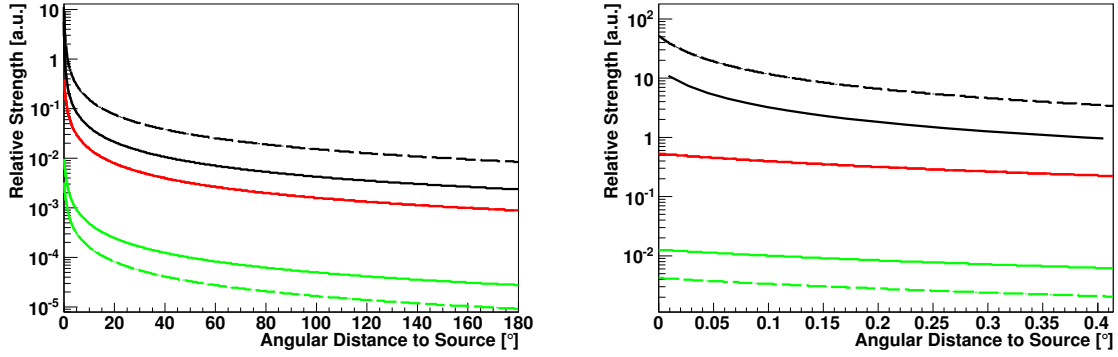


Figure 6.6: Distribution around the three sources corresponding to the Inverse model. Solid lines represent strength model A, dashed lines strength model S. Since Vir A (red) is the reference in both cases, the curve is the same. The black one represents Cen A and the green one For A. *Left Panel:* The distribution over the complete range. *Right Panel:* The distribution for small distances shows a clear difference in the shape close to the corresponding source.

6.1.5.2 Random Sources

Due to the fact that no source has been confirmed so far, we include here a model consisting of random source positions on the sky, since all are in principle equally probable. The number of considered sources depends on the overall signal to noise ratio used for the maps, for instance a contribution of 40 % in this scenario corresponds to 4 sources. In general, the number of sources is equal to the fraction of the anisotropic contribution divided by ten.

For the smearing also an alternative approach is applied. Here, a flat distribution of events around the source is considered whereupon these have to be within specified borders. Three configurations are chosen:

1. a circle with a radius of 10° ,
2. an ellipse with axis of 15° and 5° (randomly oriented) and
3. an ellipse with axis of 20° and 10° (randomly oriented).

Due to the random positions the density map resembles isotropy. Hence, we show in Fig. 6.7 only one event set consisting of 60 events from 6 ellipses plus 40 events from isotropy considering the PAO coverage. The left panel displays the case of a circle of radius 10° , while the one on the right panel gives an example for ellipses having a long axis of 15° and a short one of 5° . In the second case this stretched shape is visible for some ellipses.

6.1.6 Multipoles

These are the most prominent types in search for anisotropies in the arrival directions of cosmic rays not only at the highest energies. Particularly, those of lower order can

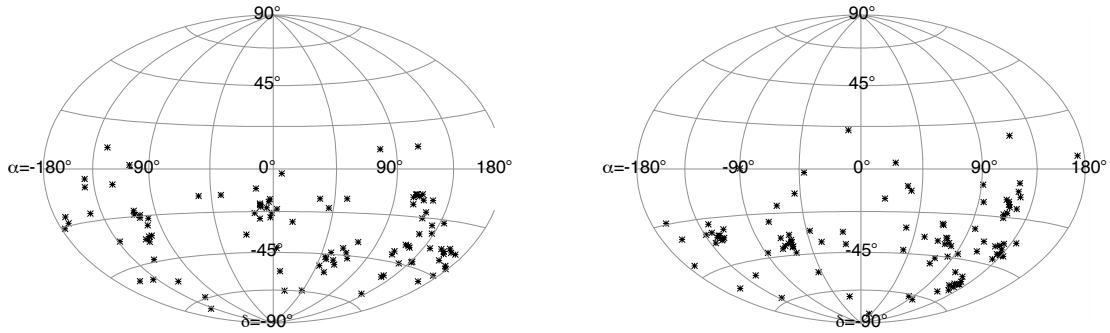


Figure 6.7: An example event set consisting of 10 events distributed in 6 ellipses each plus 40 events from the isotropic background. *Left Panel:* Six circles of radius 10° are distributed within the field of view of the PAO. *Right Panel:* Six ellipses with axis 15° and 5° are distributed within the field of view of the PAO.

be tested at an early stage of the experiment before smaller substructures become visible in the data. Therefore, only dipoles, quadrupoles and octupoles on a sphere are considered with two orientations (x, z) . Furthermore, the isotropic background usually added is neglected given that these models are large scales anisotropies distributed over the complete sky. The following equations describe the z -direction of the particular multipole:

$$\begin{aligned} \text{Dipole:} \quad f(x, y, z) &= \sqrt{\frac{3}{4\pi}} \frac{z}{r}, \\ \text{Quadrupole:} \quad f(x, y, z) &= \frac{1}{4} \sqrt{\frac{5}{\pi}} \cdot \frac{-x^2 - y^2 + 2z^2}{r^2}, \text{ and} \\ \text{Octupole:} \quad f(x, y, z) &= \frac{1}{4} \sqrt{\frac{7}{\pi}} \cdot \frac{z(2z^2 - 3x^2 - 3y^2)}{r^3}. \end{aligned}$$

with r the radius of the considered sphere, here set to one, and x, y, z the usual Cartesian coordinates. For the x -direction, just substitute the x for z and vice versa. We use these formulas as the probability density to fill the bins in the corresponding maps. In Fig. 6.8 the density distributions for all used multipoles can be seen.

6.2 Analysing Mock Maps

We apply the methods introduced in chapter 5 to the created MC sets and in the following present exemplary results based on a preselection of maps. The chosen source scenarios are chosen to be very different. In this way, a more detailed overview can be given of the capabilities of the algorithms. At this, we split the discussion according to the classification used previously in this chapter. As benchmark points we have chosen probabilities of 0.01 and 0.001. In this context, we count the number of isotropic sets yielding an even larger deviation from the isotropic expectation than the tested anisotropic MC sets, i.e. the estimators of the isotropic sets have to larger or

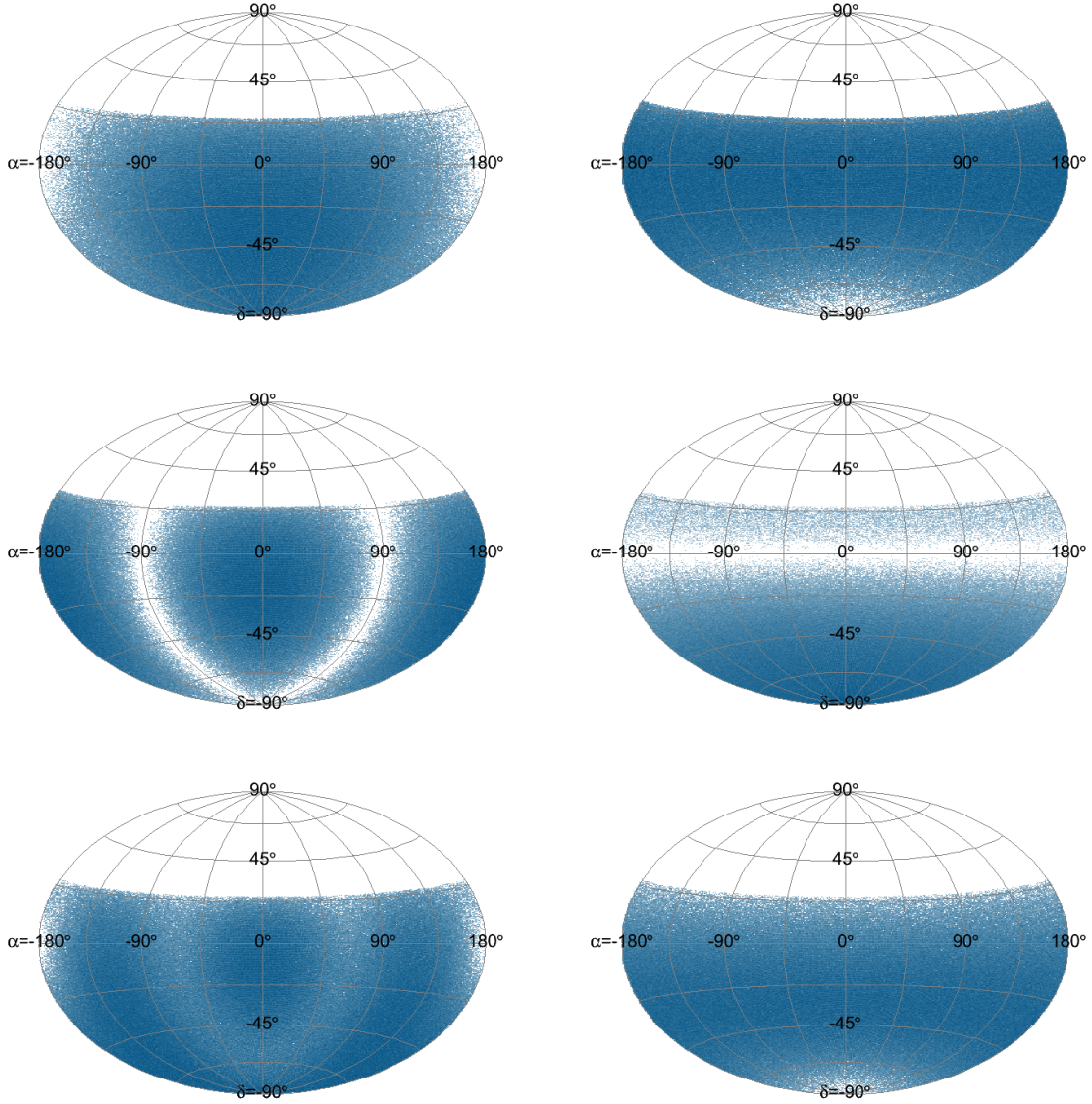


Figure 6.8: Multipoles in x (left) and z (right) direction weighted with the coverage of the PAO. *Top Panel:* Dipole. *Middle Panel:* Quadrupole. *Bottom Panel:* Octupole.

smaller than the those of the anisotropic samples depending on the algorithm. These values for P can be interpreted as errors of first kind, accepting anisotropy although the events are isotropically distributed. To estimate the efficiency and sensitivity of the autocorrelation methods, we calculate the fraction of MC sets identified as being anisotropic, therefore having smaller values of P . This corresponds to $1-\beta$, the modified error of the second kind, accepting anisotropy in the case the distribution is really anisotropic.

Regarding both variants of the Cluster Algorithm, six different maximal angular radii have been considered, from 5° up to 30° in steps of 5° . The results presented below are based on an analysis with a radius $R = 15^\circ$. In this way the best perform-

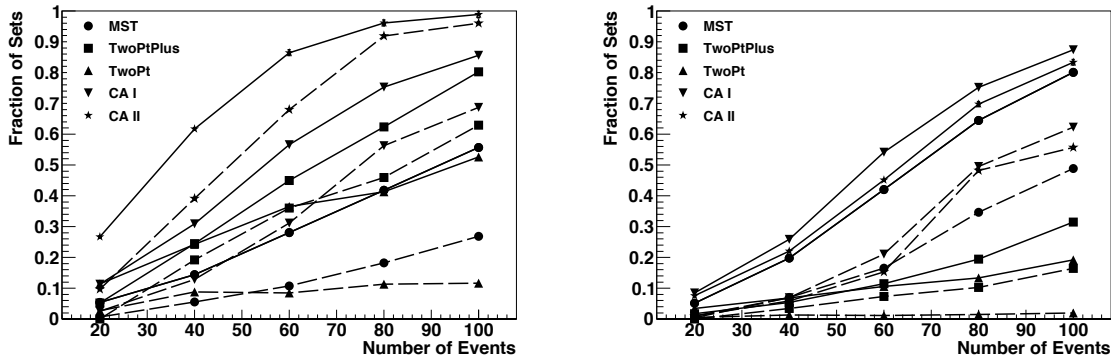


Figure 6.9: Displayed is the amount of event sets giving a probability smaller than $P=0.01$ (solid line) and $P = 0.001$ (dashed line) for all methods. *Left Panel:* MC sets based on the VCV catalogue with a smearing of 3° , the strength of a galaxy on Earth is distance weighted and the contribution of isotropy is 60 %. *Right Panel:* MC map based on the IRAS catalogue with a smearing of 6° , the strength of all galaxies on Earth is the same. The contribution of isotropy is zero.

ance is achieved based on the considered MC scenarios. A detailed discussion of the dependence of the results on R is presented in the appendix E.

In addition, we provide tables containing the results averaged over all set sizes. However, just the fraction for $P = 0.01$ is given here. In appendix F the tables for each set size are presented.

6.2.1 Analysis of catalogue-based Mock Maps

From each catalogue we have chosen one representation:

1. VCV with distance weighting, 3° smearing and 60 % isotropic background and
2. IRAS with flat weighting, 6° smearing and 0 % isotropic background.

While the first map represents the findings of the correlation analysis of the PAO, the second one is complementary. Instead of a few strong sources against an isotropic background, many weak sources form the complete flux of UHECRs.

The results are displayed in Fig. 6.9.

Regarding the sets based on the VCV catalogue, the cluster algorithm in both variants is very sensitive to this kind of anisotropy. Nearly all sets with 100 events have a probability of $P < 0.01$ and for the weighted version (CA II) this is even true for $P < 0.001$. Taking a few sources with a small spread of UHECRs around them is enough to distinguish the resulting distribution of arrival directions from isotropy. The unweighted cluster algorithm (CA I) is a little bit less effective which can be explained by the used smearing (Gaussian). This enhances CA II since it is the same function which is used to weight to the events within a cluster (see equation (5.6)). While the TwoPtPlus at least for 75 % of all sets still finds a deviation from the

isotropic expectations the TwoPt correlation function and the MST are at a level of 50 % for 100 events per set. Despite the strong background which dilutes the signal, still half of all sets are identified to be anisotropic.

For the other scenario the development regarding the sensitivity of the autocorrelation methods is different. Indeed CA I and CA II are again the best methods, however the difference to the MST is very small, of the order of a few percent. This is caused by the large number of objects contributing equally to the overall flux of UHECRs. Thus, the probability that one source emits more than one UHECR is small and identifying single sources is becoming even more challenging. However, by choosing a radius for the cluster algorithms which is larger than the smearing scale, events are combined which have not the same origin. This is the case here, since $R = 15^\circ > \sigma = 6^\circ$ which explains the results.

For the MST the good performance can be explained by the distribution of the sources which form filamentary structures. Due to the small smearing scale, these structures propagate to the distribution of the events. Since, the average distance between two objects emitting a UHECR is smaller than the average distance for isotropy a large fraction of sets deviated clearly from the Null hypothesis.

The TwoPt correlation functions are again on a lower level since many sources all with the same strength for these methods mimic an isotropic distribution. This can be seen for instance in the probability density distribution (cf. Fig. D.1, left sky plot of the bottom row) where only a small band is missing making the difference to isotropy.

In the following we give the tables containing all results regarding the analysis on catalogue-based source scenarios. To this end we have averaged over all event set sizes to obtain the fraction of sets having a probability smaller than 0.01. In Tab. 6.2 to 6.5 the average fractions are given averaged over all set sizes.

Maps with flat weighting

The general numbers presented in Tab. 6.3 for maps based on the IRAS catalogue conform the results presented above on the example map. By increasing the amount of background the sensitivity decreases being only at the level of 10 % or less for all methods for 40 % isotropic contribution. Due to the similarity to the coverage of the PAO these scenarios are hard to detect for all methods.

By reducing the number of possible sources however, the results are slightly better. Indeed more sets are identified to be anisotropic but again a background level of 60 % shows that the methods are nearly blind to this scenario. The impact of the doubling of the smearing scale is only of the order of a few percent.

Maps with distance weighting

This picture changes significantly when applying a weighting according to the distance of the objects (cf. Tab. 6.4 and 6.5). In contrast to the scenario before now only those galaxies being very close to the Earth contribute to a major amount of all events. This makes it for a cluster algorithm to detect a deviation from isotropy. The other methods give similar results. However, by increasing the background contribution they are much more affected than the CAs and suffer the loss of sensitivity. While they can still benefit from the occurring clusters, the others are weakened

Table 6.2: Presented are the fraction of sets based on the VCV with a flat weighting giving a probability smaller than 0.01 relative to isotropy averaged over all set sizes.

VCV with flat weighting										
Smearing	3°					6°				
Isotropy [%]	0	20	40	60	80	0	20	40	60	80
MST	0.67	0.40	0.14	0.04	0.01	0.61	0.31	0.10	0.03	0.01
TwoPt	0.70	0.45	0.18	0.05	0.01	0.66	0.40	0.15	0.04	0.01
TwoPtPlus	0.52	0.30	0.12	0.03	0.01	0.48	0.27	0.10	0.03	0.01
CA I	0.70	0.46	0.19	0.04	0.01	0.67	0.42	0.16	0.03	0.01
CA II	0.77	0.60	0.30	0.08	0.02	0.66	0.44	0.18	0.05	0.01

Table 6.3: Presented are the fraction of sets based on the IRAS with a flat weighting giving a probability smaller than 0.01 relative to isotropy averaged over all set sizes.

IRAS with flat weighting										
Smearing	3°					6°				
Isotropy [%]	0	20	40	60	80	0	20	40	60	80
MST	0.50	0.23	0.08	0.03	0.01	0.42	0.18	0.06	0.02	0.01
TwoPt	0.14	0.05	0.02	0.01	0.01	0.11	0.04	0.02	0.01	0.01
TwoPtPlus	0.16	0.08	0.03	0.02	0.01	0.14	0.06	0.03	0.01	0.01
CA I	0.55	0.31	0.12	0.04	0.01	0.50	0.27	0.10	0.03	0.01
CA II	0.56	0.34	0.14	0.04	0.02	0.46	0.24	0.10	0.03	0.01

since they include large angular distances which are more likely to be dominated by isotropy.

Again, the effect of doubling the smearing scale is small.

Summary

It is worth noting that despite the extend source distribution the CA methods show the best results especially in the presence of an isotropic background.

6.2.2 Analysis of single-source Mock Maps

This category is split in two separate approaches and therefore, we discuss them individually here.

Table 6.4: Presented are the fraction of sets based on the VCV with distance weighting giving a probability smaller than 0.01 relative to isotropy averaged over all set sizes.

VCV with distance weighting										
Smearing	3°					6°				
Isotropy [%]	0	20	40	60	80	0	20	40	60	80
MST	0.98	0.89	0.68	0.29	0.04	0.97	0.86	0.60	0.21	0.03
TwoPt	0.99	0.94	0.77	0.33	0.03	0.98	0.91	0.71	0.25	0.03
TwoPtPlus	0.97	0.90	0.77	0.43	0.06	0.96	0.88	0.73	0.38	0.05
CA I	0.99	0.94	0.83	0.52	0.07	0.99	0.94	0.81	0.50	0.07
CA II	1.00	0.98	0.92	0.74	0.20	0.99	0.95	0.85	0.59	0.10

Table 6.5: Presented are the fraction of sets based on the IRAS with distance weighting giving a probability smaller than 0.01 relative to isotropy averaged over all set sizes.

IRAS with distance weighting										
Smearing	3°					6°				
Isotropy [%]	0	20	40	60	80	0	20	40	60	80
MST	0.98	0.89	0.71	0.35	0.06	0.97	0.87	0.64	0.26	0.04
TwoPt	1.00	0.97	0.86	0.47	0.06	0.99	0.94	0.80	0.36	0.04
TwoPtPlus	0.97	0.91	0.78	0.46	0.08	0.95	0.86	0.70	0.34	0.05
CA I	1.00	0.96	0.88	0.64	0.14	1.00	0.96	0.87	0.64	0.14
CA II	1.00	0.99	0.96	0.85	0.41	1.00	0.98	0.92	0.75	0.24

Radio Loud Galaxies

First, two example maps are considered based on the three radio loud galaxies, Cen A, Vir A and For A which are analysed as described above:

1. Gaussian smearing, relative strength according to acceleration model S (80 % isotropic background) and
2. Inverse smearing, relative strength according to acceleration model A (40 % isotropic background).

In the left panel of Fig. 6.10, displaying the result of the first map, the CA methods again show the best performance due to the few sources contributing a major amount of events. The S model has the characteristic that Cen A is responsible for more than 90 % of all events. This is, however, not enough to result in a significant deviation compared to isotropy concerning the MST, which ends up with a fraction of 20 % and below. Both TwoPt functions give fractions in between 40 and 60 %.

By changing the smearing function and the acceleration model the sensitivity of the CA methods drop by about a factor of two (cf. right panel of Fig. 6.10). Hence, if in nature the deflection mechanisms leads to an Inverse smearing the possibility to recognize anisotropy with these methods decreases significantly. It is worth noting

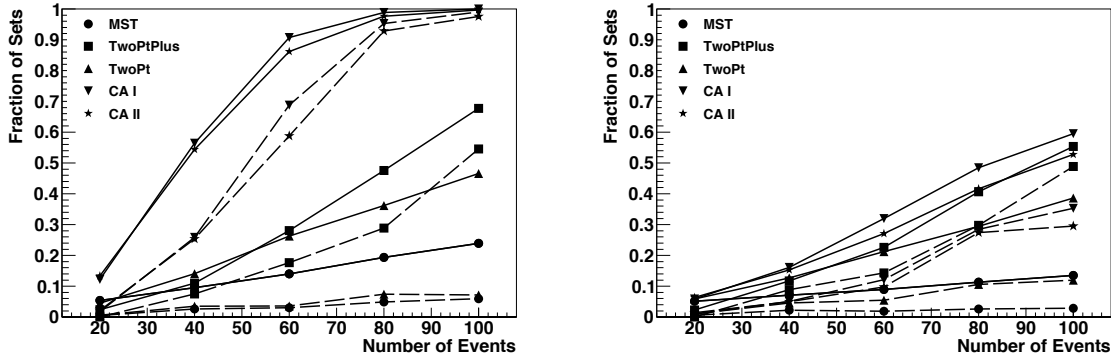


Figure 6.10: Displayed is the amount of event sets giving a probability smaller than $P=0.01$ (solid line) and $P = 0.001$ (dashed line) for all methods. The only contributing objects are Cen A, Vir A and For A. *Left Panel:* MC sets analysed are based on acceleration model S with a Gaussian smearing. The contribution of isotropy is 80 %. *Right Panel:* MC sets analysed are based on acceleration model A with a smearing according to the Inverse model. The contribution of isotropy is 40 %.

that this happens although the isotropic background has been reduced to 40 %. The other methods remain nearly unchanged in their performance while the TwoPt correlation functions are then comparable in sensitivity to the CAs.

Including also the other simulated realisations, the results are found in Tab. 6.6 and 6.7 for the Gaussian and Inverse smearing, respectively, which are discussed in the following.

Maps with a Gaussian smearing

For the smallest amount of isotropy all methods show an excellent performance. More than 90 % of all sets are identified to be anisotropic. However, the behaviour regarding an increase of background is different. While the CAs are always above the 60 - 70 % level, the MST ends up with values close to 10 % and the TwoPt correlation functions between 20 % and 30 %.

This is results is expected due to the similarities of scenario and design of the cluster algorithms.

Maps with an Inverse smearing

As already discussed above the change in the smearing function affects all methods drastically. Especially, the CAs lose sensitivity although only very few sources contribute to the overall flux of cosmic rays. The reason for this lies in the density distribution. Although 68 % of the events emitted by one of the three sources lie within the smearing scales σ (given in equation (6.8)) around the corresponding sources, 99 % of the events spread over the complete sphere and are not contained within circles of 3σ around the sources. For the Gaussian smearing, however, this relation is valid. Hence, a clustering is disfavoured by this model. In addition, the broad distribution on the complete sphere only leads to a small alteration of the distribution of arrival directions as compared to isotropy which in turn causes a drop in sensitivity for the other methods. Here, It is worth noting that a single strong

Table 6.6: Presented are the fraction of sets based on the radio loud galaxies with a Gaussian smearing giving a probability smaller than 0.01 relative to isotropy averaged over all set sizes.

Radio loud galaxies with a Gaussian smearing						
strength model	A			S		
Isotropy [%]	40	60	80	40	60	80
MST	0.93	0.61	0.11	0.96	0.70	0.14
TwoPt	1.00	0.89	0.19	1.00	0.93	0.26
TwoPtPlus	0.98	0.84	0.28	0.99	0.86	0.31
CA I	1.00	0.97	0.66	1.00	0.99	0.72
CA II	1.00	0.96	0.64	1.00	0.98	0.70

source as in model S is easier detected than two weaker sources in model A.

Table 6.7: Presented are the fraction of sets based on the radio loud galaxies with an Inverse smearing giving a probability smaller than 0.01 relative to isotropy averaged over all set sizes.

Radio loud galaxies with an Inverse smearing						
strength model	A			S		
Isotropy [%]	40	60	80	40	60	80
MST	0.09	0.04	0.02	0.13	0.05	0.02
TwoPt	0.22	0.07	0.02	0.32	0.10	0.02
TwoPtPlus	0.27	0.08	0.02	0.30	0.09	0.02
CA I	0.32	0.11	0.02	0.44	0.17	0.03
CA II	0.29	0.10	0.02	0.39	0.15	0.03

Ellipses

The two example maps analysed based on the ellipse model are:

1. four ellipses containing 10 % of all events each with a radius of 10° and
2. four ellipses containing 10 % of all events each with a large axis of 15° and a small one of 5° .

For each scenario, isotropy accounts for the other 60 % of events.

Due to the model description one can assume that CA I and CA II should be particularly sensitive and indeed they show again the best performance over the different set sizes (cf. Fig. 6.11). Already for a small number of events a large amount of sets is recognized to be anisotropic, not only for $P=0.01$ but also for $P=0.001$. Furthermore, for small event sets CA I is more sensitive than CA II. However, for 80 events per set both are at the maximum.

The MST and the TwoPt functions show also a very good performance reaching

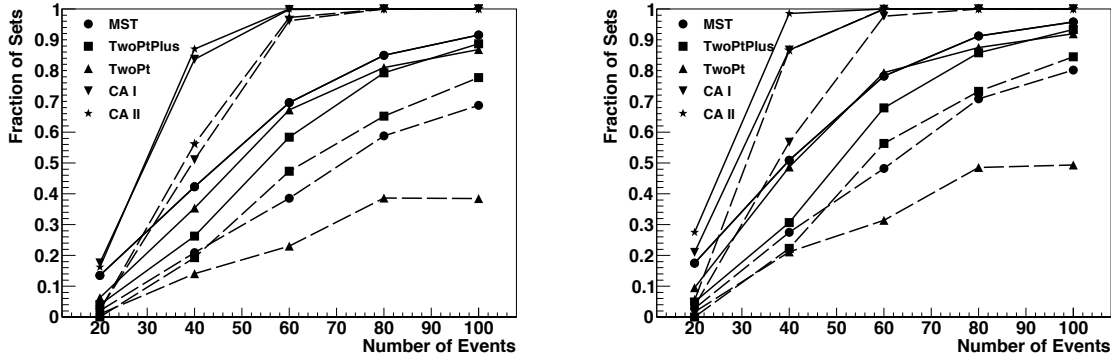


Figure 6.11: Displayed is the amount of event sets giving a probability smaller than $P=0.01$ (solid line) and $P = 0.001$ (dashed line) for all methods. In both cases four ellipses are considered each accounting for 10 % of all events. The other 60 % are contributed by isotropy. *Left Panel:* The ellipses are circles with a radius of 10° . *Right Panel:* The ellipses have a large axis of 15° and a small one of 5° .

Table 6.8: Presented are the fraction of sets based on the Ellipse giving a probability smaller than 0.01 relative to isotropy averaged over all set sizes.

Ellipses with dimensions									
Axes	$10^\circ, 10^\circ$			$15^\circ, 5^\circ$			$20^\circ, 10^\circ$		
Isotropy [%]	40	60	80	40	60	80	40	60	80
MST	0.85	0.60	0.15	0.88	0.67	0.18	0.83	0.52	0.12
TwoPt	0.76	0.55	0.16	0.82	0.63	0.20	0.78	0.54	0.15
TwoPtPlus	0.68	0.51	0.21	0.72	0.57	0.25	0.61	0.41	0.15
CA I	0.89	0.80	0.58	0.90	0.82	0.61	0.85	0.74	0.46
CA II	0.87	0.81	0.62	0.91	0.85	0.72	0.81	0.71	0.43

nearly a fraction of 100 % for 100 events. Among these three methods the differences are very small of the order of a few percent.

The fraction for all simulated maps can be found in Tab. 6.8. All methods are similar in their performance with advantages on the side of the CAs especially in the presence of a large isotropic background where it still can identify up to 70 % of all anisotropic event sets. The other methods again are weakened due to the increasing contribution of isotropic events to the estimator. The resulting fraction then is of the order of 10 - 20 %.

6.2.3 Analysis of Multipoles

At last we consider two example sets based on multipoles with the corresponding results presented in Fig. 6.12:

- a dipole in x direction and
- an octupole in z direction.

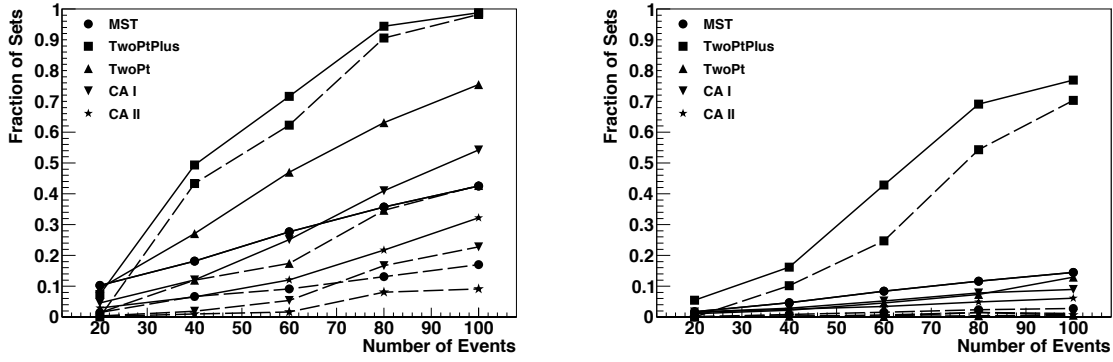


Figure 6.12: Displayed is the amount of event sets giving a probability smaller than $P=0.01$ (solid line) and $P = 0.001$ (dashed line) for all methods. *Left Panel:* The dipole is in x direction *Right Panel:* The octupole is in z direction.

While for the dipole (see left panel of Fig. 6.12) some methods are performing very well (TwoPt and TwoPtPlus have a fraction $> 70\%$), the other three are at the level of 50% and below. This has been expected for the CAs due to their design. An important factor is the number of maxima in the density distribution and how distinct they are.

Furthermore, the analysed sets based on an octupole show a very different behaviour regarding all methods. The TwoPtPlus is least affected and still on a reasonable level regarding the fraction of identified sets as compared to the other methods which are below the 10% level. In general, the difference between this octupole and isotropy is too small which can be seen by comparing the underlying density distribution with the coverage of the PAO.

Results for all multipole models are given in the summary Tab. 6.9. The first thing to notice is that no method shows an overall good performance. At least for one map the sensitivity is very low, at the percent level.

Furthermore, the quadrupole, especially the one in z direction, seems to produce event sets which can be identified easier as being anisotropic. This is caused by the typical structure. The strong maxima in the field of view of the PAO enhance pairs, triplets and so forth on small angular scales. Studying the second row of Fig. 6.8 one can also understand why the MST is weaker for the x than for the z direction. In z direction most of the events are contained in a region below the equator, only a small amount lies above. This leads to a smaller length of the MST due to the smaller available area on which the events can be distributed. Similar considerations explain why the TwoPt correlation functions are sensitive.

6.3 Summary and Conclusions

In this chapter many MC models have been presented which were the basis for a study aiming to find the best autocorrelation method among the ones introduced in the previous chapter. Many different scenarios have been covered regarding number of sources, their strength, distribution, smearing function and scale and so forth.

Table 6.9: Presented are the fractions of sets based on multipoles giving a probability smaller than 0.01 relative to isotropy averaged over all set sizes.

Multipole Direction	Dipole		Quadrupole		Octupole	
	x	z	x	z	x	z
MST	0.27	0.01	0.48	0.96	0.05	0.08
TwoPt	0.44	0.85	0.92	0.99	0.02	0.01
TwoPtPlus	0.64	0.37	0.82	0.67	0.04	0.42
CA I	0.27	0.01	0.55	0.88	0.06	0.05
CA II	0.15	0.01	0.38	0.78	0.04	0.04

These were grouped depending on the considered type of anisotropy into three classes, namely catalogue based, few sources and multipoles. The few sources model is again divided into two subclasses characterised by different source positions and smearing.

In the end, one is interested in the average fraction of sets having probabilities smaller than 0.01 relative to isotropy. To avoid a bias towards one type of anisotropy we first calculated the mean fraction for each class of anisotropy. In total, for the catalogue based model 40 different maps have been analysed, 6 maps based on multipoles, 12 using the radio loud galaxies and and 9 with an elliptical shape for the sources. The corresponding results per map as well as the average over all maps can be found in Tab. 6.10. For the second case we have averaged over all models, one time including multipoles and the other time excluding this scenario. This selection can be motivated by the obtained results by analysing these maps. All methods are able to identify at least one kind of multipole. However, no method shows an overall good performance. In addition, many more orientations are possible for the multipoles making it hard to estimate the behaviour of all methods. Therefore, we exclude multipoles and check which method gives then the best result. In both cases the resulting numbers are very similar and the order regarding the sensitivity is nearly unaltered.

The weakest method which finds only 40 % of all anisotropic sets is the MST. One reason for this is the influence of isotropically distributed background events. Since their average distance towards another events is larger and leads therefore in turn to a larger value for the estimator. A possible way to reduce this factor is to define a maximum edge length, i.e. the MST is calculated and afterwards all edges are removed being longer than a predefined value. This corresponds to an application of the MST presented by Harari et al. in [70]. However, the maximum length has to be motivated by theory or estimated by exploring the data.

On the second place we find the TwoPt and TwoPtPlus correlation functions yielding a slightly better sensitivity than the MST. As we have seen, the behaviour of these two is very similar for all presented maps. A clear difference could only be found for certain multipoles whereas these fluctuations are characteristic for all methods. Hence, by using the TwoPtPlus the gain in sensitivity is in general small. The reason for the difference in the performance of the both methods might be an effect of the coverage of the PAO. Further comparisons considering the complete sphere could

Table 6.10: Fraction of sets with $P < 0.01$ relative to isotropy averaged over all MC maps and set sizes

Maps	MST	TwoPt	TwoPtPlus	CA I	CA II
Catalogues	0.38	0.39	0.37	0.46	0.51
Radio loud galaxies	0.32	0.42	0.42	0.54	0.52
Ellipses	0.53	0.51	0.46	0.74	0.75
Multipoles	0.31	0.55	0.50	0.30	0.23
all maps (w/ MP)	0.38	0.47	0.43	0.51	0.50
all maps (w/o MP)	0.41	0.44	0.41	0.58	0.59

test if their performance is the same or what causes the differences. It is worth noting that there might exist an optimal choice for the reference coordinate system for the TwoPtPlus leading to better results. A good candidate for this is the super galactic coordinate system where the equator contains many AGN close to Earth among them Cen A.

Regarding the CA methods, both show a better performance than the other three which is strengthened by excluding multipoles. The difference of the results between those two versions of the CA is just about 1%. Hence, the effect caused by the weighting within the cluster is only minimal in the end by using these simulated models as a reference.

Thus, for the final analysis of the data of the PAO we choose a CA. To decide which of the two realisations is applied to the event set two arguments are discussed. The first one is to use a method and an analysis which are as simple as possible. In addition, we can take current anisotropy results into account, i.e. the correlation of events of the PAO with AGN. By assuming that the true sources are distributed according to the matter distribution in our cosmological neighbourhood, the performance of the CAs on this kind of MC maps should have more weight in the process of deciding. From Tab. 6.2 to 6.5 one can see that the CA II is more sensitive to this type of anisotropy leading to the decision to use it for the analysis of the data of the PAO presented in the next chapter.

7. Application to Data measured with the PAO

In the last chapter it has been shown that the weighted Cluster Algorithm (CAII) with a radius $R = 15^\circ$ works best on the chosen MC models. Therefore, in the following it is applied to the high energy data set based on measurements of the surface detector of the PAO. First, we describe the cuts constraining the data set to events of high quality. They are the same which have been so far used for all publications of the PAO collaboration regarding anisotropy studies at energies larger than 50 EeV, e.g. the correlation of UHECRs with AGN. Moreover, to compare the autocorrelation method with the catalogue dependent study we use this published analysis as a benchmark for our method, i.e. we compare the resulting probability for the identical data set.

However, this publication considers only events which have been detected prior to the end of 2009. Hence, we extend the data set as far as possible. In this context, the two independent event reconstruction algorithms existing within the collaboration of the PAO are compared.

In addition, we consider all events with energies larger than 3 EeV. Thus, it can be checked whether the magnetic field (galactic as well as intergalactic) suppresses clustering on small scales.

Finally, we move on from autocorrelation to point source searches. We try to identify a possible source candidate by looking at the positions of the found clusters. The vicinity of the Centaurus A is of particular interest here.

7.1 Cut Parameters

The first cut removes bad periods from the data set. At these times, the data acquisition was unstable and thus, a proper event reconstructions cannot be guaranteed. In the end it is recommended to exclude three periods for anisotropy studies at the highest energies which are listed in Tab. 7.1:

Furthermore, each event has to pass the complete trigger chain described in chapter 4.2.1. Only those fulfilling a 6T5 or a 5T5 trigger are allowed, i.e. the tank with the strongest signal has to be surrounded by at least 5 operating detector stations. In the following, only those events are considered which are well reconstructed and have a zenith angle smaller than 60° . Events with a larger value for θ are reconstructed in a different way as mentioned before and until now the energy scales are not matched. Therefore, an integration in the analysis could lead to questionable results. However,

Table 7.1: Bad periods excluded from the used data set.

Begin					End				
Day	Month	Year	Hour	Minute	Day	Month	Year	Hour	Minute
26	3	2005	18	15	6	4	2005	23	56
22	4	2005	0	0	2	5	2005	23	57
17	4	2009	0	0	6	6	2009	23	59

when this issue is resolved, it is worth to include these events due to a significant increase of statistics of roughly 25 %.

Another important step is the energy cut. If the cut value is too low the signal will overwhelmed by the isotropic contribution depending on the strength of the signal. If it is too high the amount of events might be too small due to the steep energy spectrum. One way to deal with this is to scan the data sets for the optimal radius yielding the smallest probability for an agreement with the isotropic expectation. In the context of the AGN correlation study an optimal value for the energy cut has been found in this way. The best results were obtained for a lower energy limit of $E_{cut} = 55$ EeV. However, a different analysis will usually end up with a different value. The mentioned cuts are summarized in Tab. 7.2:

Table 7.2: Standard Quality cuts applied to data set

Trigger:	5T5 6T5
Reconstruction	OK
Zenith Angle:	$\theta < 60^\circ$

7.2 High Energy Data Sets

Within the PAO two independent reconstruction frameworks exist (CDAS and Offline). Each one yields a file containing all reconstructed events, which are called Herald and Observer, respectively. These are used in the following. The first one is chosen to be the basis for all anisotropy analyses at the highest energies, for instance in the AGN correlation publications. However, we consider both to have an estimate regarding systematic uncertainties caused by the reconstruction scheme.

As a start, we analyse the same data sets as used in the last publication on anisotropy studies published by the PAO. Afterwards, the number of events is updated until the 31st of October 2010 with the same cut parameters, since the list of bad periods is missing for the time after. Both reconstructions are compared and the resulting differences are discussed.

For the following analyses the considered data set is compared to 10,000 isotropic sets of the same size, simulated according to the coverage of the PAO. It must be noted that for the isotropic events only the arrival direction in equatorial coordinates is simulated, i.e. neither an air shower nor a detector simulation is included. The smearing of the original values according to the detector resolution can be neglected, since a smeared isotropic distribution is equal to the non-smeared one.

Regarding the data of the PAO, these events have passed through the complete reconstruction chain and hence, the original values are already smeared.

7.2.1 Comparison with the result of the AGN correlation study

The corresponding data set consists of 69 UHECRs fulfilling the cut mentioned above including the cut on the energy. A correlation analysis of these events with AGN of the VCV catalogue lead to a probability of 10^{-3} that the cosmic rays are isotropically distributed on the sky.

Applying the Cluster Algorithm II with the maximal radius of 15° to this data set we find an excess of clusters of the size seven, nine and ten (see top left panel of Fig. 7.1) compared to 10,000 isotropically distributed sets. However, by looking at the spread which is based on these sets, one can see that the data is well contained in the expectation of isotropy. The corresponding weights marked as stars can be found in the bottom left panel together with the expected density of isotropic clusters. The weights are in agreement with the expectation from isotropy. The estimator of the CA II is the Cluster Sum which is displayed in the bottom right panel of Fig. 7.1. It exhibits no clear discrepancy between the isotropic expectation based on 10,000 sets (black histogram) and the value (red line) for the data. In the end, this yields a probability of $P = 0.04$ for a Cluster Sum of 39.49. Thus, with the CA II no deviation from the isotropic expectation could be found.

7.2.2 Current Data Set

The 55 EeV lower energy cut was appropriate for the AGN correlation analysis which was found by a scan. Thus, a different study will in general have an optimal value for E_{cut} differing from this one, i.e. the optimal energy cut yields the smallest probability for an agreement between data and isotropy.

Instead of performing also a scan, we followed here the approach presented in [51]. In this context, the event set is increased in steps of 20 up to 100 events starting with the 20 of the highest energy. The used parameters and the corresponding Cluster Sum and probability can be found in Tab. 7.3.

Table 7.3: For different set sizes corresponding to different lower energy cuts, E_{cut} , the Cluster Sum (CS) and corresponding probability (P) are given for the Herald.

N	20	40	60	80	100
E_{cut} [EeV]	76.61	66.32	60.67	55.44	52.12
CS	4.35	8.94	23.80	42.34	74.86
P	0.10	0.54	0.30	0.26	0.06

For all considered event sets, no deviation from isotropy is found. Although, for the 20 highest energetic events the probability is close to the $P = 0.01$. This could hint towards a weak clustering signal in the data of the PAO. Future data will help to answer the question if this is just a fluctuation or a true trend towards anisotropy.

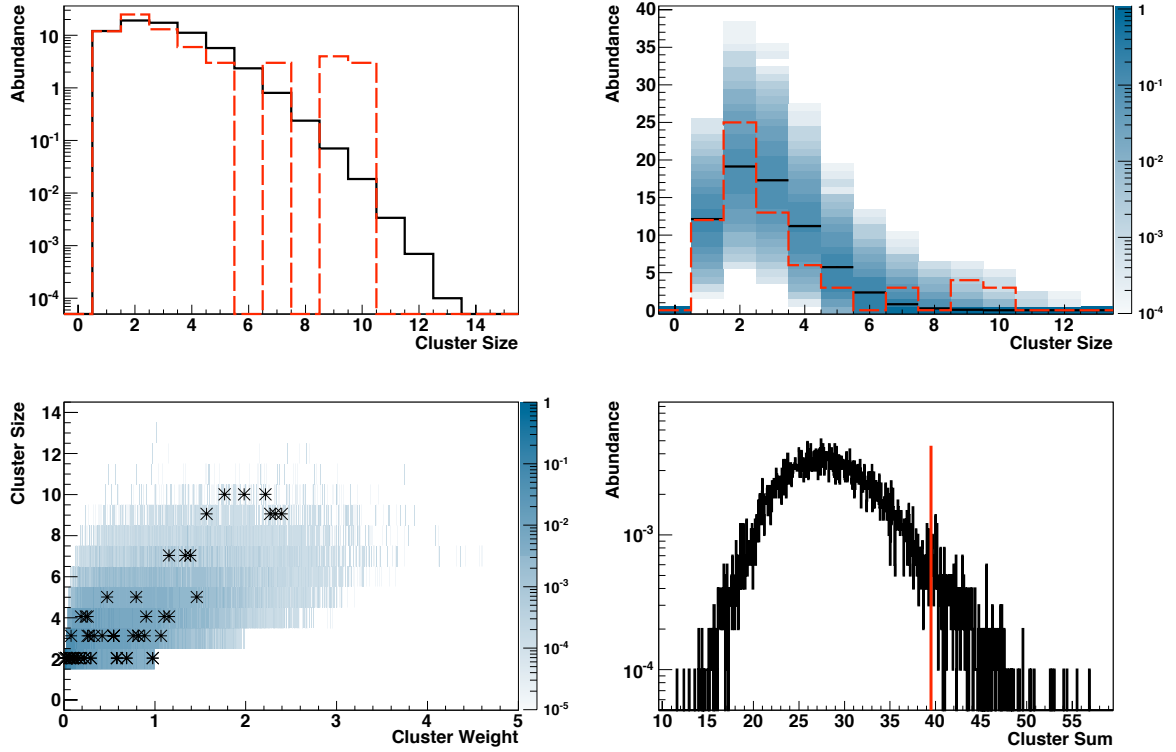


Figure 7.1: Presented are the results of the CA II applied to the 69 published high energy events of the PAO which are compared to the results of 10,000 isotropic sets of the same set size. *Top Left Panel:* The isotropic expectation of the Cluster Size is shown in black and in red dashed the data with logarithmised y-axis. *Top Right Panel:* The mean of the isotropic expectation of the Cluster Size is drawn in black, in red dashed the data and in addition, in blue shade the density of the expected spread for isotropic sets based on simulations. *Bottom Left Panel:* The blue shade represents the density of the weights expected for isotropic sets and the stars mark the calculated weights of the data set. *Bottom Right Panel:* The black distribution corresponds to the Cluster Sum of the isotropic sets and the red line to the Cluster Sum of the 69 events detected by the PAO leading to a probability of 0.04 that this value is in agreement with isotropy.

7.2.2.1 Comparison between Herald and Observer

In the following, we compare the events sets provided by the two reconstruction schemes. To this end, we consider the 100 of the highest energy events out of Herald and Observer. A comparison on an event-by-event basis yields that only 88 events have the same identification number while the remaining 12 are completely different. This can be explained for example by a differing energy assignment to a certain event (see below).

The differences in energy ($\Delta E = (E(\text{Herald}) - E(\text{Observer}))/E(\text{Herald})$) and arrival direction reconstruction ($\Delta\Omega$) of shared events having an identical time stamp are shown in Fig.7.2.

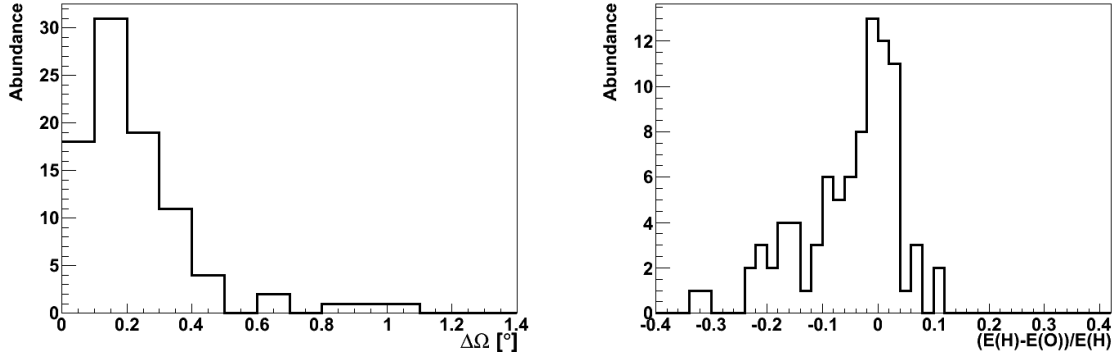


Figure 7.2: Comparison of the two reconstruction algorithms in energy and arrival direction for the 88 events of the highest energy. *Left Panel:* The mean difference between Herald and Observer in the arrival direction is 0.2° , i.e. well below the angular resolution of the detector. *Right Panel:* Observer contains energies which are on average 5 % smaller than those found in the Herald whereas the RMS is about 9 %.

The mean of both distributions is below the resolution of the detector, which is 1° on the reconstructed arrival direction and 10 % on energy (pure statistical error for energies of $E = 10^{19}$ eV):

$$\begin{aligned}\Delta\Omega &= 0.2^\circ, \\ \Delta E &= 5\%.\end{aligned}$$

The resulting Cluster Sum for the Observer and Herald are displayed in Fig. 7.3 including the isotropic expectation. All calculated values for the Cluster Sum are in good agreement with the Null hypothesis. The corresponding probabilities are found in Fig. 7.4 as well as in Tab. 7.4 with further parameters, namely event size and the corresponding energy.

Table 7.4: For different set sizes corresponding to different lower energy cuts, E_{cut} , the Cluster Sum (CS) and corresponding probability (P) are given for the Observer.

N	20	40	60	80	100
E_{cut} [EeV]	79.45	67.80	61.42	56.08	51.03
CS	3.67	10.92	23.31	45.14	68.45
P	0.18	0.30	0.33	0.16	0.17

7.3 Low Energy Data Set

As discussed before one expects at lower energies that no anisotropy should be present due to the stronger influence of the magnetic field, galactic as well as intergalactic. Hence, we tested this by analysing events with an energy larger than 3 EeV

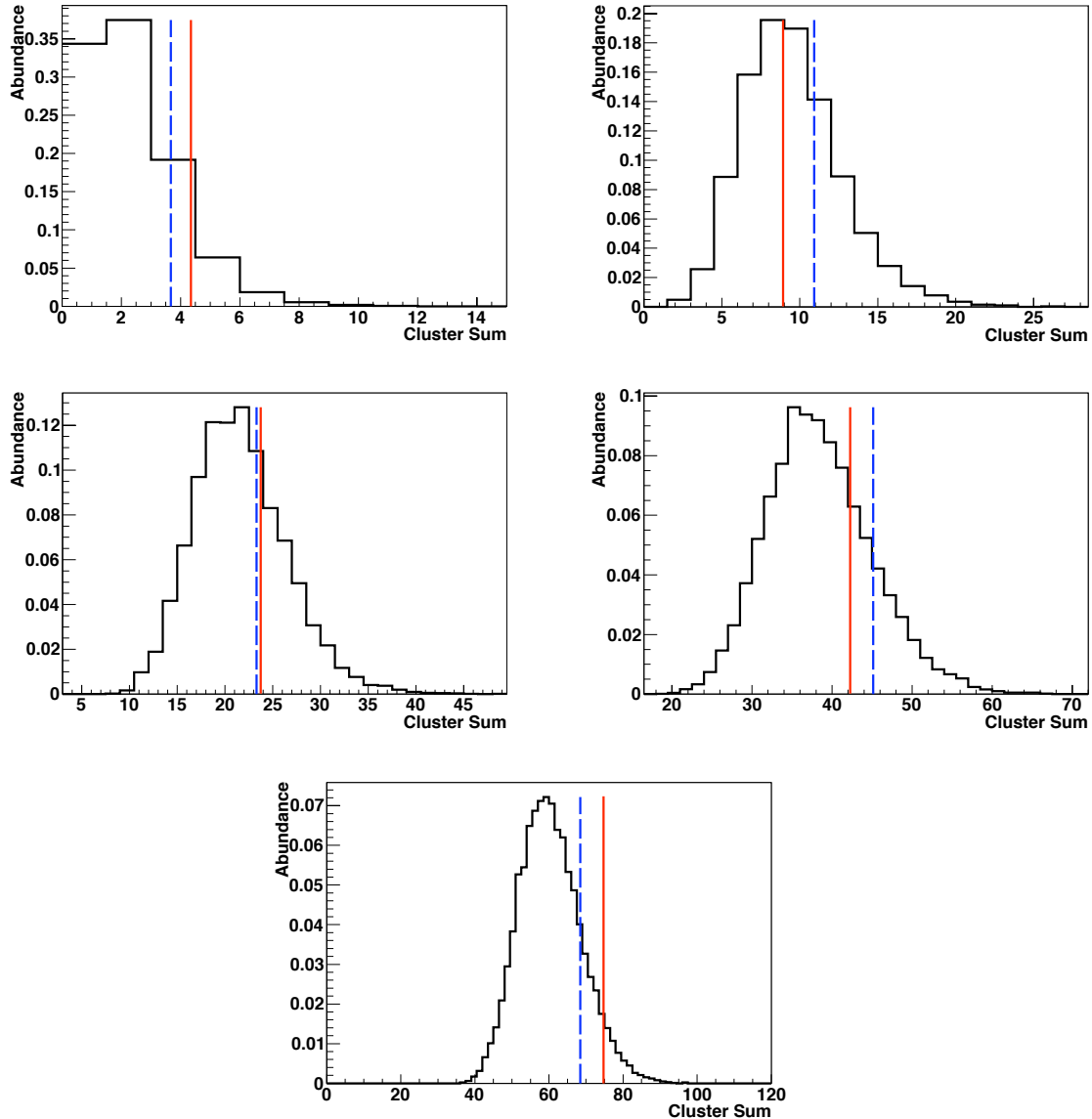


Figure 7.3: Displayed is the Cluster Sum of 10,000 isotropic sets in black for sets sizes of 20, 40, 60, 80 and 100, from left to right and top to bottom. The red solid line corresponds to the Herald and the blue dashed to the Observer.

surviving all cuts mentioned before. In addition, a different reference file for bad periods is in use including even more time periods up to the 31st of October 2010 than the one given above. However, this can introduce a false large scale anisotropy together with the geomagnetic field and the weather effect (as discussed in [48]). These local effects can cause an anisotropy in the arrival directions in equatorial coordinates due to their influence on the energy reconstruction. For a detailed study, we refer to the Ph.D. thesis of M. Grigat.

This cross-check here is intended to take a first look at low energy data. Future studies including the relevant effects will therefore change the below presented results.

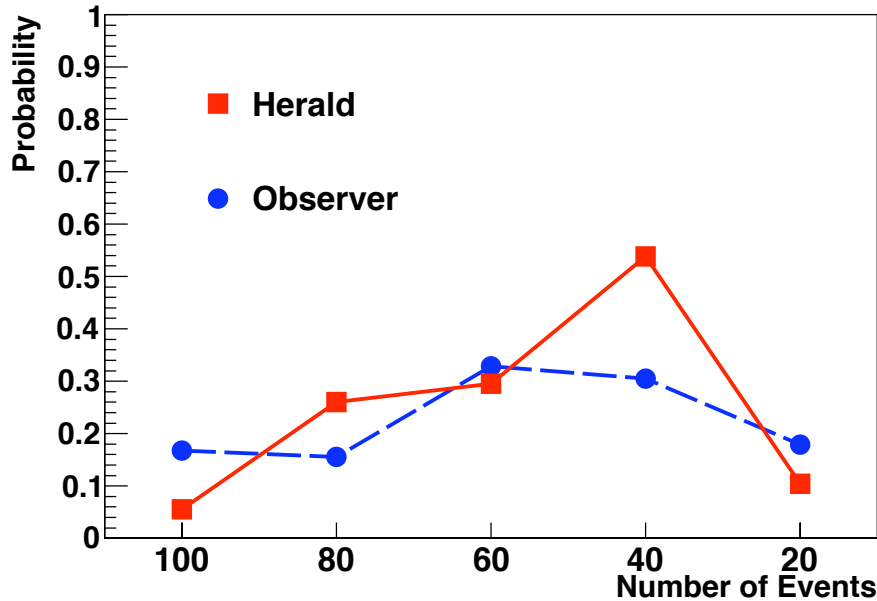


Figure 7.4: Probability of the data sets based on Herald (red, solid line) and Observer (blue dashed line) for different sizes. The lines are only drawn to guide the eye. For both schemes no significant deviation from isotropy is found.

For this analysis the Herald file is used, which yields more than 44,800 events. These are split in sets containing 100 events each analysed with the CA II. In the left panel of Fig. 7.5 the abundance of the Cluster Size normalised to one set is presented in red for the data and in black for isotropy. The latter is again based on 10,000 sets. A difference is visible in particular for high multiplicities, i.e. clusters containing more than 10 events. This becomes even clearer by looking at the right panel of the same Fig. where the ratio of the two histograms is shown.

However, the corresponding distributions of the Cluster Sum for isotropy (black) and data (red) (cf. Fig. 7.6) show a similar behaviour. We applied a χ^2 -test to estimate the comparability between these two distributions which resulted in:

$$\begin{aligned}\chi^2 &= 51.24, \\ \text{Ndof} &= 42 \text{ and} \\ \text{Probability} &= 15\%.\end{aligned}$$

Thus, no deviation from isotropy is found so far. However, a future study using the Cluster Algorithm II, which accounts for local effects, could reveal an underlying anisotropy.

7.4 Point Source Search

Although, the autocorrelation is a proper approach to search for anisotropies as we have seen before, in the end one would like to identify the sources or at least interesting regions. In this context, another advantage of the cluster algorithm

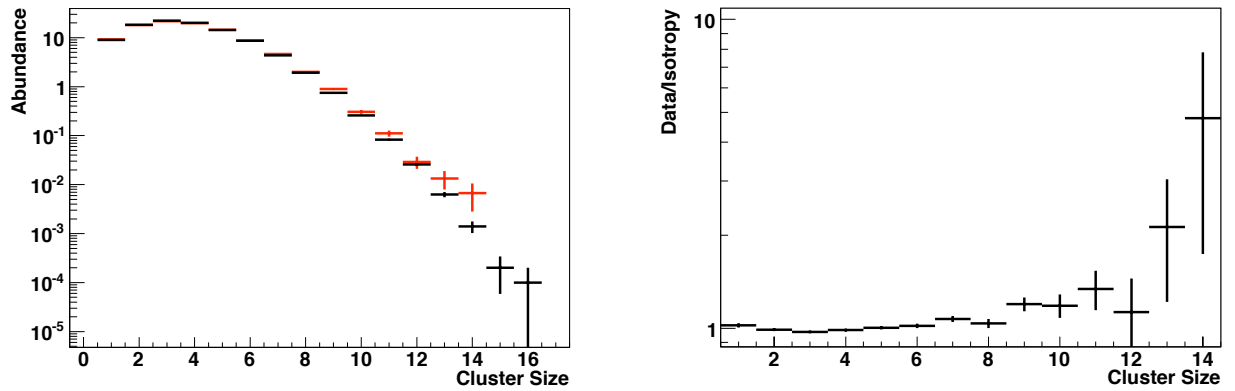


Figure 7.5: *Left Panel:* Cluster Size distribution is displayed for the low energy data set where the isotropic model corresponds to the black points and the data to the red points. Both event sets are normalised to one set. *Right Panel:* The corresponding ratio of both histograms, data divided by isotropy.

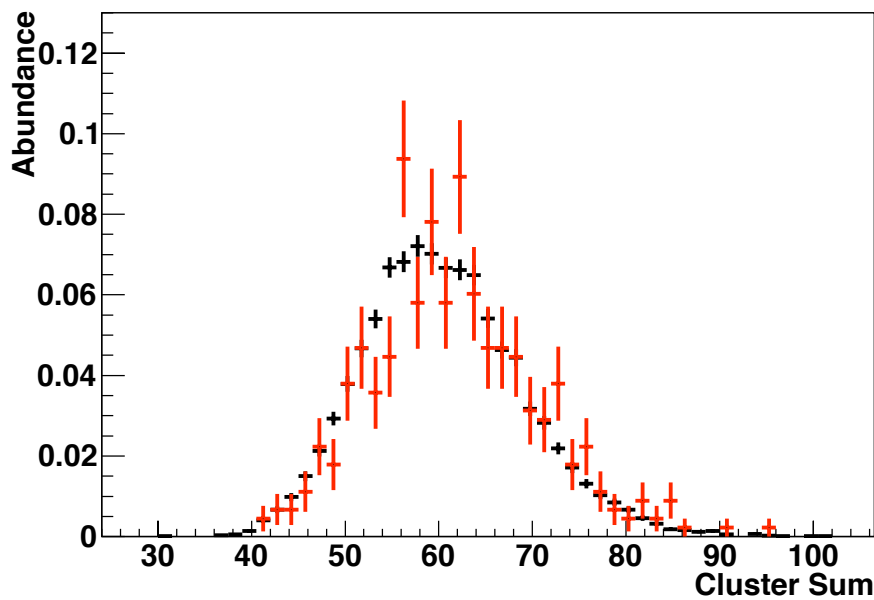


Figure 7.6: In red the distribution of the Cluster Sum based on 448 low energy data sets of the PAO containing 100 events each. Black symbols indicate the expected distribution for isotropy. A comparison of both with a χ^2 test yields a probability of 15 %.

should be pointed out. It keeps the information about the position of the found clusters which other methods cannot provide, e.g. the TwoPt.

In chapter 3 it has been discussed that especially one object might be responsible for many UHECRs detected at Earth. Cen A has drawn much attention due to its physical properties (distance to Earth, extended radio lobes) and the high amount of events in its vicinity. Therefore, we study in the following the region around Cen A. A first a posteriori analysis was performed in [25] (cf. Fig. 3.7) focusing on the angular distance Ψ_{CenA} of the events towards Cen A.

This has, however, the disadvantage of including all events, i.e. also those isolated and hence being most likely background.

Therefore, we include only clusters found by the CA II, their position and weight. We analysed the 100 events of highest energy provided by the Herald. The resulting histogram filled with the weight of the cluster at the corresponding distance (red, dashed) can be seen in the top panel of Fig. 7.7 as well as the mean expectation of 10,000 isotropic sets in black. A peak is visible around 40° corresponding to the south pole where the exposure is the largest, leading in turn to a higher probability of clusters of different sizes. For the largest distance the histograms have empty bins since this represents regions on the sky which are not covered by the standard field of view of the PAO.

While only a comparison with the mean distribution would lead to a clear statement towards a deviation from isotropy we include the spread based on all sets leading to the blue shaded area. This is again the density distribution for isotropy similar to the one presented in the context of the introduction of the TwoPt correlation function. To compare both distributions, we also adapted the estimator based on the negative Log-likelihood and calculated its value for the 10,000 isotropic sets and the data set (cf. bottom panel of Fig. 7.7). Only two sets give a larger value for the estimator leading to a probability $P = 2 \cdot 10^{-4}$.

However, this is still no proof for Cen A to be the source of UHECRs. This very region on the sky is very dense on objects and thus, all of them could be responsible for the UHECRs. Furthermore, a strong directional magnetic field within or outside our galaxy might focus the particles into that region on the sky.

It is again worth noting that this is an a posteriori analysis. To establish, however, a true anisotropy in the direction of Cen A a prescription is needed for instance.

7.5 Summary

In this chapter the CA II has been applied to high energy as well as to the low energy data set of the PAO. Both analysis yield an agreement with the isotropic expectation. Although, the probability obtained by analysing the data set containing 100 events might hint towards a clustering on small scales. It must be noted that for the latter one no corrections were applied to the data set accounting for the influence of weather and geomagnetic field effect. Hence, an analysis including these changes could still find an anisotropy.

In addition, another advantage of the CA II has been used. In contrast to other autocorrelation methods, the CA II keeps the information where the clusters are located. Thus, the field of application can be extended to point source searches.

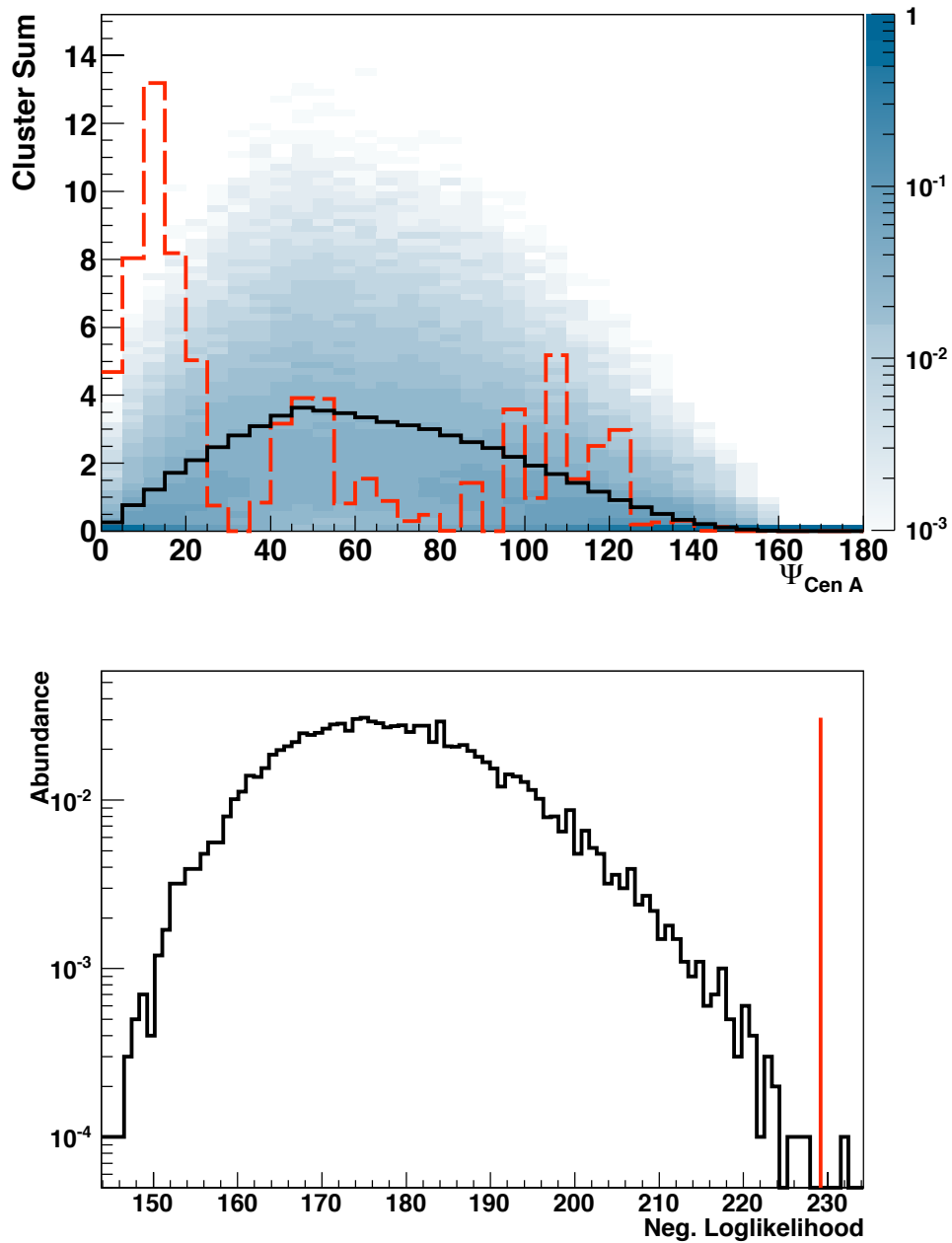


Figure 7.7: *Top Panel:* Distance distribution of Clusters to Cen A. The histograms are filled with the weight of the corresponding clusters. The mean expectation for isotropy is shown in black and in blue shade the underlying density distribution. The data is again displayed as the red and dashed line. *Bottom Panel:* In black the resulting distribution of the negative Log-likelihood estimator for isotropy. The red line corresponds to the result based on the 100 highest energetic events leading to a probability of $2 \cdot 10^{-4}$.

Here, we have focused on Cen A, a very prominent source candidate for UHECRs. However, this analysis was a posteriori and the calculated probability of $2 \cdot 10^{-4}$ has to be confirmed by future analyses.

8. Conclusion & Outlook

In this work, recent results of the PAO have been presented including an anisotropy study of the arrival direction of cosmic rays at the highest energies. In this context, a correlation analysis with nearby AGN has been performed leading to a first establishment of anisotropy on the southern hemisphere.

In this work a different approach has been presented focusing on autocorrelation methods as a tool to search for a deviation from the isotropic expectation. Three different algorithms have been introduced relying only on the angular distance between two events $\Omega_{i,j}$. Each one of them is specialised for a certain kind of anisotropy. The first one, a TwoPt correlation function, is a standard method and represents the reference for all others. A first step towards a new method was to extend this known one. While the original representation compares the distribution of solid angles between all events of isotropy and the data set, the TwoPtPlus uses instead the distribution of the two angular distances within the two angles in a reference spherical coordinate system. This, however, leads to a dependence of the method on the choice of it. In general, an infinite number of coordinate systems is available which can be more or less motivated. e.g. the equatorial system which is the reference coordinate system of the PAO and hence, the natural choice.

These methods consider the relations of all events independent of their angular distance which might lead to a dilution of the signal by isotropy. To compensate this effect we have adapted a concept out of graph theory, the Minimum Spanning Tree (MST). Here only the closest events contribute forming a graph connecting all events according to Prim's algorithm without loops. In this way especially filamentary structures become visible on the sky map.

The third type of method, the cluster algorithm (CA), reduces not only the scale of angular distances contributing to the estimator but also the number of events. A maximal radius R is defined leading to pairs, triplets and so forth of events which are closer to each other than R . Isolated events, however, are ignored since they are most likely from isotropy. The weight of each cluster can be computed in two different ways. While in the first (CA I), the angular distances within a cluster are irrelevant leading to a weight equal to the number of events inside, in the second case (CA II) the weight is calculated by means of a Gaussian with $3\sigma = R$. In this way also those clusters are suppressed where $\Omega_{i,j}$ is close to R which is again very typical for isotropy. The sum over all weights is then the Cluster sum the estimator of both realisations.

In this thesis the mentioned methods have been tested on many types of anisotropies. To this end, several source scenarios have been simulated based on catalogues, few isolated sources as well as on a small selection of large scale multipoles. We

have varied the number of events, the contribution of the background (isotropy), the smearing scale and smearing function. It is worth noting that this selection covered only a small amount of possible theories and models.

A method has found a deviation from isotropy when the probability of a data set was smaller than $P = 0.01$. In this context, we compared the methods among themselves to find the one being most sensitive to the considered anisotropies.

In general, one could see for all methods that an increase in statistics leads to a higher sensitivity. Regarding the catalogue-based as well as the few isolated sources, the CA performs best especially for the second type which is expected due to its design. However, by changing the smearing function from a Gaussian towards an behaviour proportional to $1/\Omega$ the sensitivity is reduced drastically.

For multipoles the results are ambivalent. No method gave the best results for all types of simulated multipoles while the TwoPtPlus and TwoPt were more sensitive than the others.

Averaging over all scenarios and set sizes the CAs gave in the end the best results. The difference between both versions was just of the order of 1 %. Furthermore, the assumption that the TwoPtPlus might be better in its performance compared to the TwoPt could not be confirmed. However, further tests with a different reference coordinate system than the equatorial coordinate system will possibly change this result. A good choice might be the supergalactic coordinate system. Of course, this depends strongly on the type of anisotropy which is realised in nature.

The final decision was made based on the better performance of CA II on the catalogue based scenarios since these are encouraged by recent results of the PAO.

Finally, we applied CA II to the current data set of the PAO. In a first step the events of the latest publication regarding the AGN correlation study are analysed to compare the resulting probabilities. The CA II gave for the 69 events a probability of 0.04 which might hint towards an underlying anisotropy but so far is still compatible with the Null hypothesis.

Afterwards, the data set was updated until the 31st of October 2010 and again analysed. The 20, 40, ..., 100 events of highest energies were processed showing for all subsets an agreement with the isotropic expectations while for the 100 events (≈ 52 EeV) the lowest probability of a few percent was reached at a similar lower energy threshold than for the catalogue-dependent study. By increasing the statistics even more, this might lead to an establishment of anisotropy in the arrival directions of UHECRs by means of autocorrelation methods.

Furthermore, the energy threshold has been lowered to 3 EeV where the PAO reaches its full acceptance. The total amount of events was split in more than 400 sets consisting of 100 events. The resulting distribution of the estimator of CA II was compared again to the one based on the isotropic expectation by means of a χ^2 test. In this first simple analysis no deviation could be found. A more detailed study including local effects (geomagnetic field, weather corrections) could change this result, because these effects cannot only cause an anisotropy but also obscure it.

In the last step another advantage of the cluster algorithm has been used. It is not only suited to detect a global anisotropy but the cluster algorithm can also be used in a search for sources of UHECRs since it keeps the position information of the

found clusters. Here, only one possible source has been considered, Cen A. By many theories it has been argued that this object is an excellent candidate to accelerate UHECRs due to its vicinity and large jets. To this end, the cluster distributions around Cen A has been compared to isotropy with a negative Loglikelihood yielding a probability of $2 \cdot 10^{-4}$. It must be noted that this is an a posteriori analysis and future data is needed to confirm this number.

In sum, many aspects of anisotropies in the arrival directions of UHECRs detected with the PAO have been examined:

1. Autocorrelation study at the highest energies;
2. Autocorrelation study of events above 3 EeV;
3. Point Source Search.

While in the first two no clear signal could be found, in the third a strong deviation has been discovered which, however, needs more statistics to establish this result.

In the future when more events are available above 50 EeV then the Cluster Algorithm might be able not only to give hints towards anisotropy but depending on the signal strength to determine the position of the sources of UHECR. A publication regarding the performance of the Cluster Algorithm will follow soon.

A. Acronyms

A.1 List of abbreviations

AERA	Auger Engineering Radio Array
AGN	Active Galactic Nuclei
AIRES	AIRshower Extended Simulations
AMIGA	Auger Muon-detectors and Infill for the Ground Array
CA I	Cluster Algorithm (unweighted)
CA II	Cluster Algorithm (weighted)
CDAS	Central Data Acquisition System
CMB	Cosmic Microwave Background
CMS	Compact Muon Solenoid
CORSIKA	COsmic Ray SIMulation for KAskade
CS	Cluster Sum, estimator of CA I and CA II
EAS	Extensive Air Shower
FADC	Flash Analog Digital Converter
FD	Fluorescence Detector
FLUKA	FLUktuierende KAskade
GPS	Global Positioning System
GRB	Gamma-Ray Burst
GZK-cutoff	Greisen-Zatsepin-Kuzmin-cutoff
HEAT	High Elevation Auger Telescope
LEP	Large Electron Positron storage ring
LDF	Lateral Distribution Function
LHC	Large Hadron Collider
MSSM	Minimal SuperSymmetric Model
MST	Minimum Spanning Tree
PBH	Primordial Black-Hole
PMT	Photomultiplier tube
QGS-Jet	Quark-Gluon-String model with Jets
SD	Surface Detector
SHDM	Super Heavy Dark Matter
TwoPt	2-point-correlation function using the relative angular distance
TwoPtPlus	2-point-correlation function using the angular distance in each angle of the co
UHECR	Ultra-High Energy Cosmic Ray
VEM	Vertical Equivalent Muon

B. Tilt Monitoring System

B.1 Calibration Constants (Tilt-Monitoring Setup)

For both axes of the sensors a polynomial of I th order is used as calibration function with the corresponding coefficients a_i and the measured inclination α_i .

$$\text{CalibrationCurve: } f(a_i; \alpha) = \sum_{i=1}^n a_i \cdot \alpha^i. \quad (\text{B.1})$$

We give here the calibration constants for the inclination sensor 143. In Tab. B.1 the coefficients are given for the x-axis and in Tab. B.2 for the y-axis, respectively. Those for all other sensors can be found in the TILT-Monitoring manual.

Table B.1: Calibration constants for the x-axis of the inclination sensors.

Sensor No.	a_1	a_2	a_3	a_4
143	0.987639	6.51049e-05	1.25916e-4	9.54672e-07
Sensor No.	a_5	a_6	a_7	
143	-6.029e-07	-1.75768e-09	1.38511e-09	

Table B.2: Calibration constants for the y-axis of the inclination sensors.

Sensor No.	a_1	a_2	a_3	a_4
143	0.992402	-0.000824292	-7.79218e-06	1.44276e-05
Sensor No.	a_5	a_6	a_7	
143	2.0425e-06	-1.00341e-07	-6.95069e-09	

B.2 Software Screenshots

The LabView program is arranged in four tabs displayed in Fig. B.1. The first one (on the top left) shows the input variables required to start the program. Regarding the distance sensors, the serial number of the μ -box as well as the firmware and dll-version are needed. For the inclination sensors only their number, name and the corresponding COM ports have to be put in. Afterwards, the program runs by itself and needs in general no maintenance. The next tab on top handles the communication with the monitoring database, i.e. the correct database is set and the filling is turned on and off. Both screenshots on the bottom show the interactive plotting of the program which can be turned on if required. While for the distance sensors a relative plotting was chosen, absolute measurements are displayed for the inclination sensors.

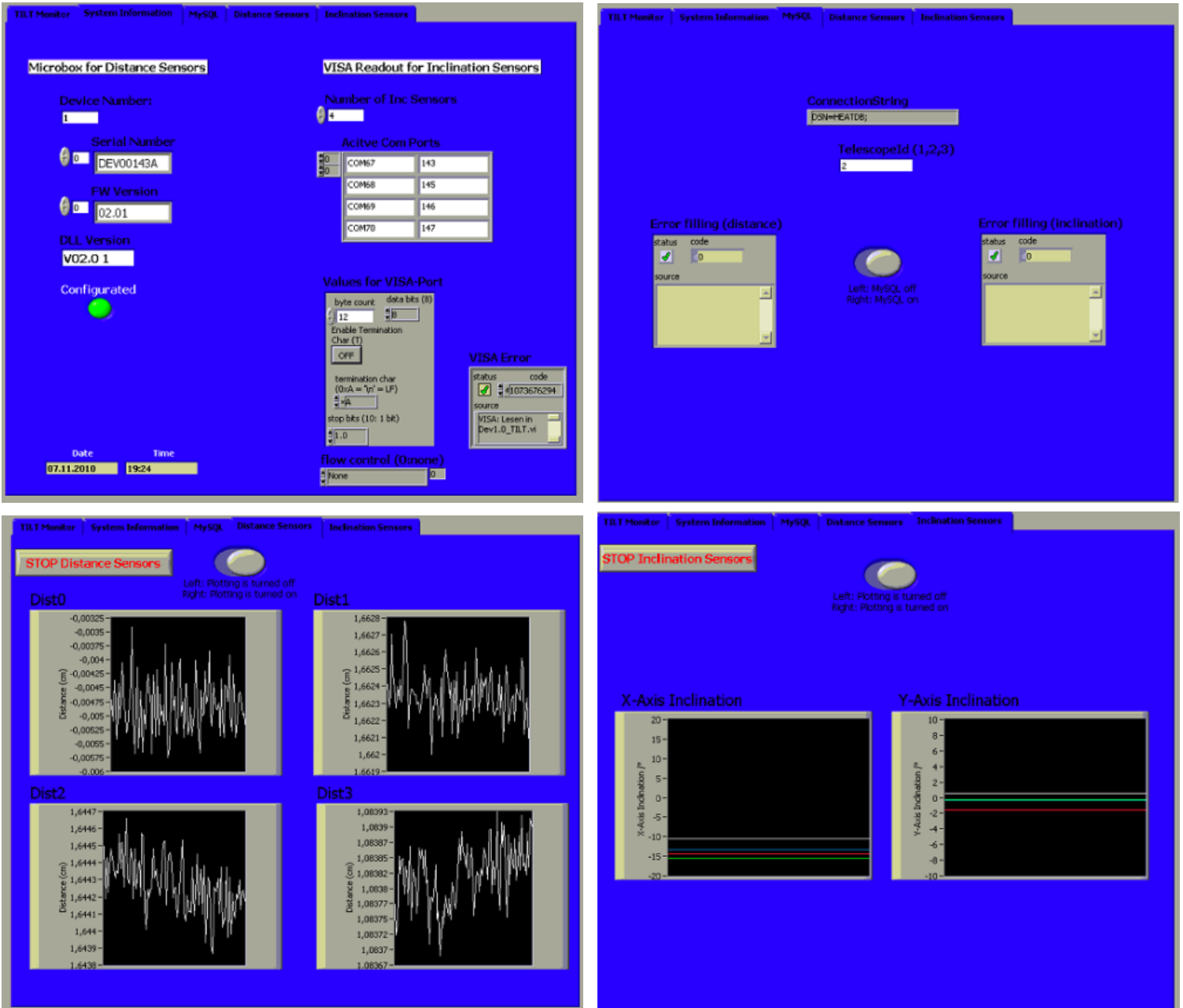


Figure B.1: Displayed are the tabs of the monitoring software visible on the screen: Top left: The system start up check. Top right: MySQL error tab. Bottom left: Interactive Distance Measurements. Bottom right: Interactive Inclination Measurements

C. Significance Calculation

C.1 Significance Calculation in the case of combined tests

In chapter 5 a realisation of the 2pt-Correlation Function is presented where two statistical tests, namely the negative loglikelihood (LL) and the Kolmogorov-Smirnov (KS) tests, are combined to evaluate the agreement between a Monte Carlo data set and the H_0 -hypothesis. In the following we discuss the effect on the resulting probability and how to handle the outcome.

In contrast to the example data set in chapter 5, we compare two independent isotropic sets. The assumption is to obtain a flat probability distribution in the end since both sets have the same origin. Note that, we consider isotropy on the complete sphere without the PAO coverage.

First, each method is used separately on the test data sets giving a distribution for LL and δ_{KS} as well as the corresponding probability. For the combination of both tests, we demand that the reference sets have to exceed the values of both estimators (cf. Fig. 5.5). The probability is then calculated by counting the number of reference sets fulfilling this requirement. In contrast to a 1-dimensional test, several points in the KS-LL plane can give the same small probability particularly for large values of LL and/or δ_{KS} . In the top panel of Fig. C.1 a sketch of the probability calculation is shown. Considering the coloured stars as the results of three isotropic sets which are compared to the reference isotropic distribution in black. Each of the three sets has a resulting probability of zero. On the other hand, the possibility to obtain a large probability is reduced. Thus, the resulting probability is biased towards small values as one can see in the left bottom panel of Fig. C.1.

To correct for this, we reuse the wrong probability distribution as a new estimator, i.e. the fraction of reference sets providing an even smaller value for P gives the true probability. In the right bottom panel the resulting distribution including a polynomial fit of 0th order is shown. The probability that this function describes the behaviour is 32 % with a $\chi^2/N_{dof} = 21.42/19$.

This approach can be adapted for any number of combined tests, e.g. the 2pt-Correlation Function using two angles, the TwoPtPlus.

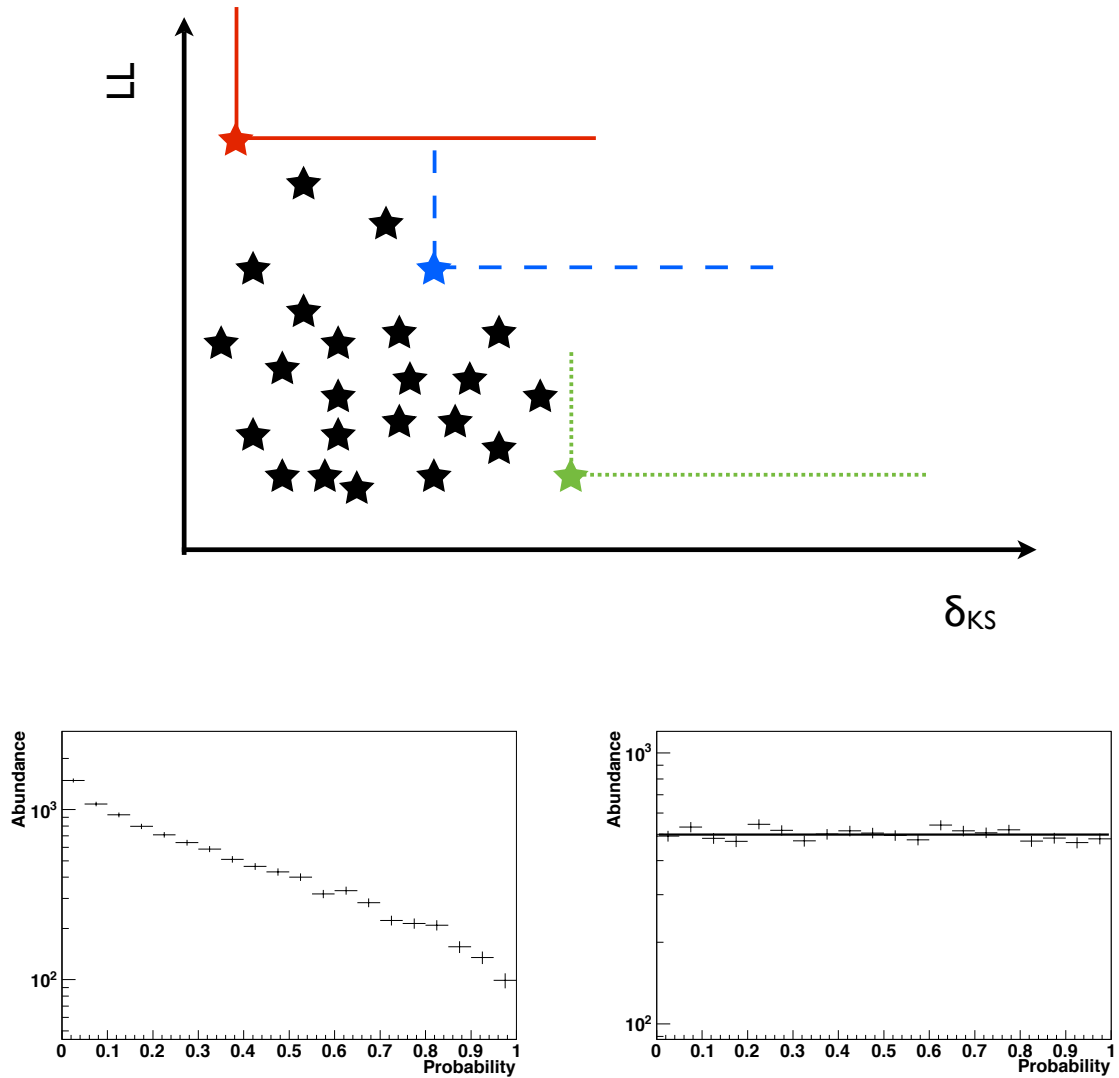


Figure C.1: *Top Panel:* A sketch of the LL-KS plane is shown with the black stars representing the isotropic reference distribution. The coloured stars represent three isotropic test sets. Each one yields a probability of zero since no reference set exceeds the values of both estimators. Thus, by combining these two tests the probability is biased towards smaller values. *Left Bottom Panel:* Uncorrected Probability for the Combination of KS and LL test for the 2pt correlation function. *Right Bottom Panel:* The true probability distribution for the combination of both tests. It is flat as expected for the case that two samples are compared having the same origin. The polynomial fit gives a $\chi^2/N_{dof} = 21.42/19$ and a probability of 32 %.

D. Monte Carlo Maps

D.1 Produced Monte Carlo Maps

In chapter 6 a procedure has been discussed to generate Monte Carlo Sets based on different source models. In this context a few maps have been presented as examples. During this study, however, several more maps have been produced as mentioned in that chapter.

Here, we give a more detailed view on those maps. The multipoles and the model based on ellipses are excluded. Since all maps regarding multipoles are already shown in the corresponding chapter and for the second one all maps would look like isotropy due to the randomly distributed positions of the sources.

Maps based on catalogues

For both catalogues (VCV and IRAS) the same parameter sets have been used, i.e. a smearing of either 3° or 6° and either a weighting of the sources according to their distance or a flat distribution. In addition, the contribution of the isotropic background is varied between 0 % and 80 %.

In Fig. D.1 the resulting IRAS density maps are displayed. As expected, for the weighted case with the smaller smearing angle, few sources are still visible even in the presence of a large background. By enlarging the smearing angle as well as increasing the number of possible sources, the substructures vanish and the maps are nearly indistinguishable from pure isotropy by eye.

A similar behaviour is also observable for VCV MC maps (see Fig. D.2). However, due to the smaller number of objects (roughly a factor of 10) they are still recognisable despite a large amount of background.

Maps based on radio loud galaxies

In addition to the catalogue based MC maps, we have also introduced a model based on three radio loud galaxies, Cen A, Vir A and For A. In this context, two different smearing algorithms, Gaussian and Inverse, and relative source strength scenarios (A and S) have been applied to the three objects. All combinations for isotropic contributions of 40 % and 80 % can be found in Fig. D.3.

As one can expect from the smearing functions, the sources are more extended in the Gaussian case (top four hammer projections) than in the Inverse (lower four hammer projections). Furthermore, the difference between S (first and third row) and A (second and fourth row) is only visible for the Gaussian smearing.

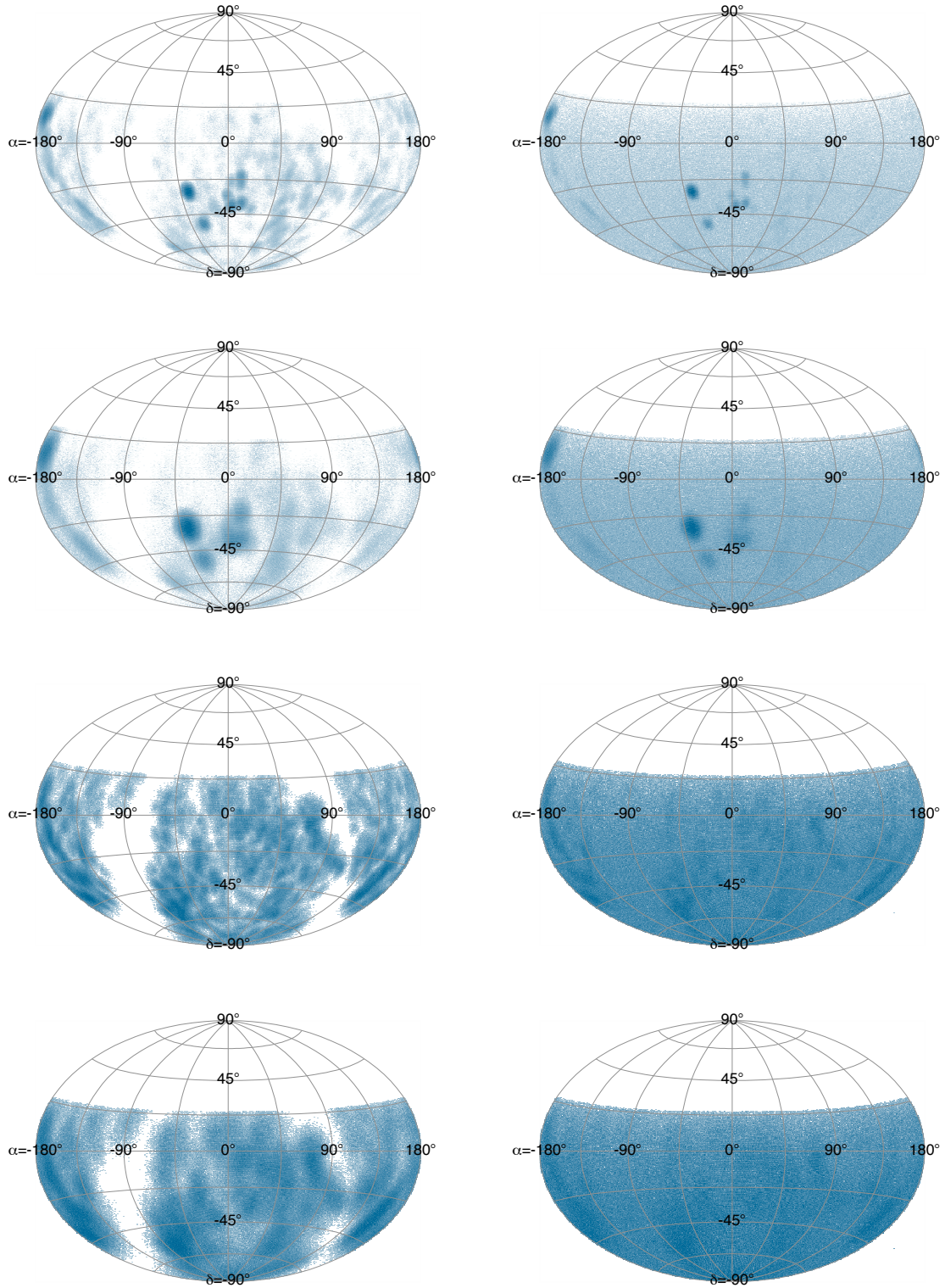


Figure D.1: MC maps based on the IRAS catalogue using different parameter sets. On the left the pure contribution is shown, while on the right 80 % of the events are contributed by isotropy. *Top Panel:* The smearing angle is 3° and the sources are weighted accordingly to their distance. *Upper Middle Panel:* The smearing angle is 6° and the sources are weighted accordingly to their distance. *Lower Middle Panel:* The smearing angle is 3° and the sources all have the same strength on Earth. *Bottom Panel:* The smearing angle is 6° and the sources all have the same strength on Earth.

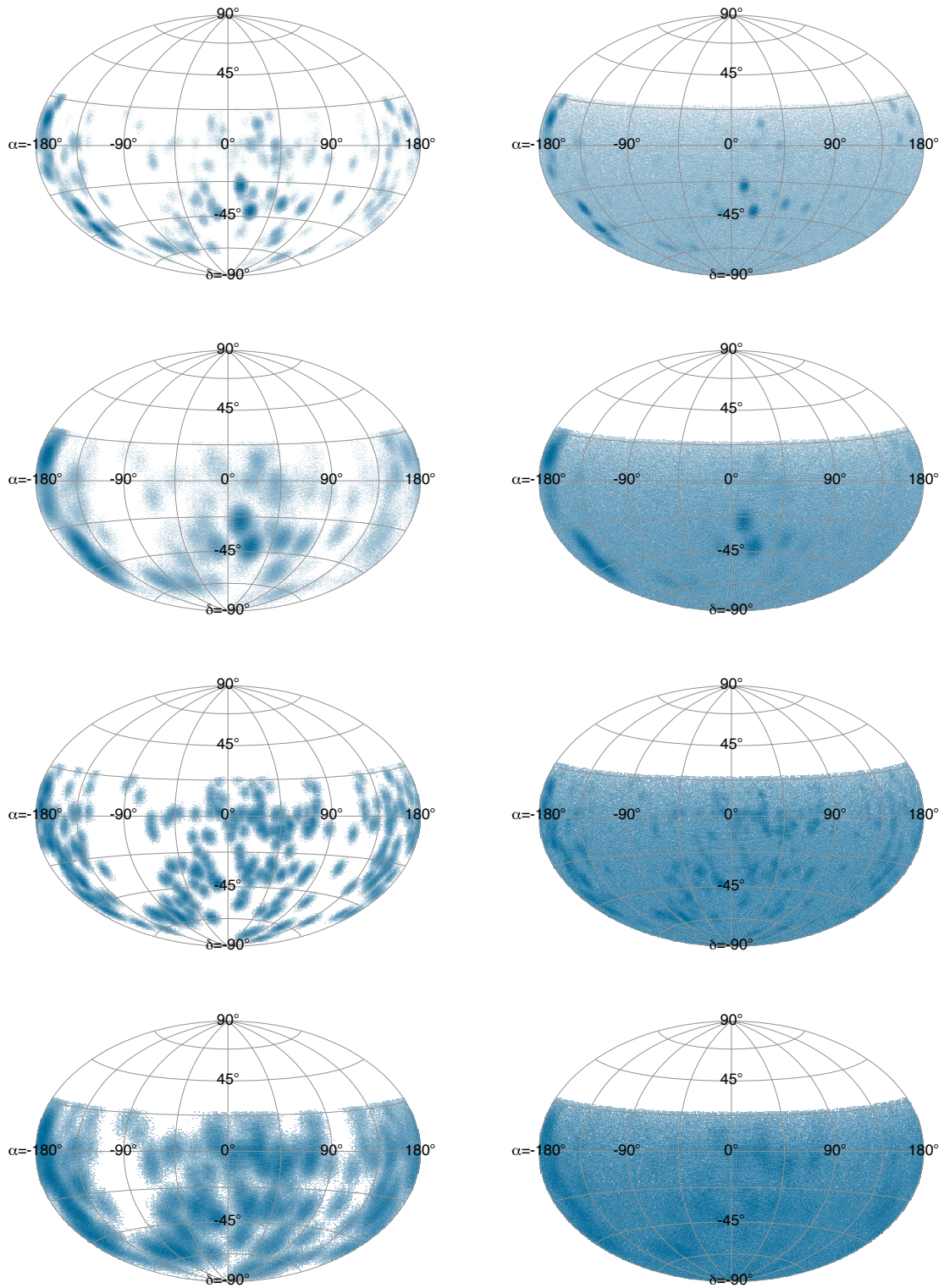


Figure D.2: MC maps based on the VCV catalogue using different parameter sets. On the left the pure contribution is shown, while on the right 80 % of the events are contributed by isotropy. *Top Panel:* The smearing angle is 3° and the sources are weighted accordingly to their distance. *Upper Middle Panel:* The smearing angle is 6° and the sources are weighted accordingly to their distance. *Lower Middle Panel:* The smearing angle is 3° and the sources all have the same strength on Earth. *Bottom Panel:* The smearing angle is 6° and the sources all have the same strength on Earth.

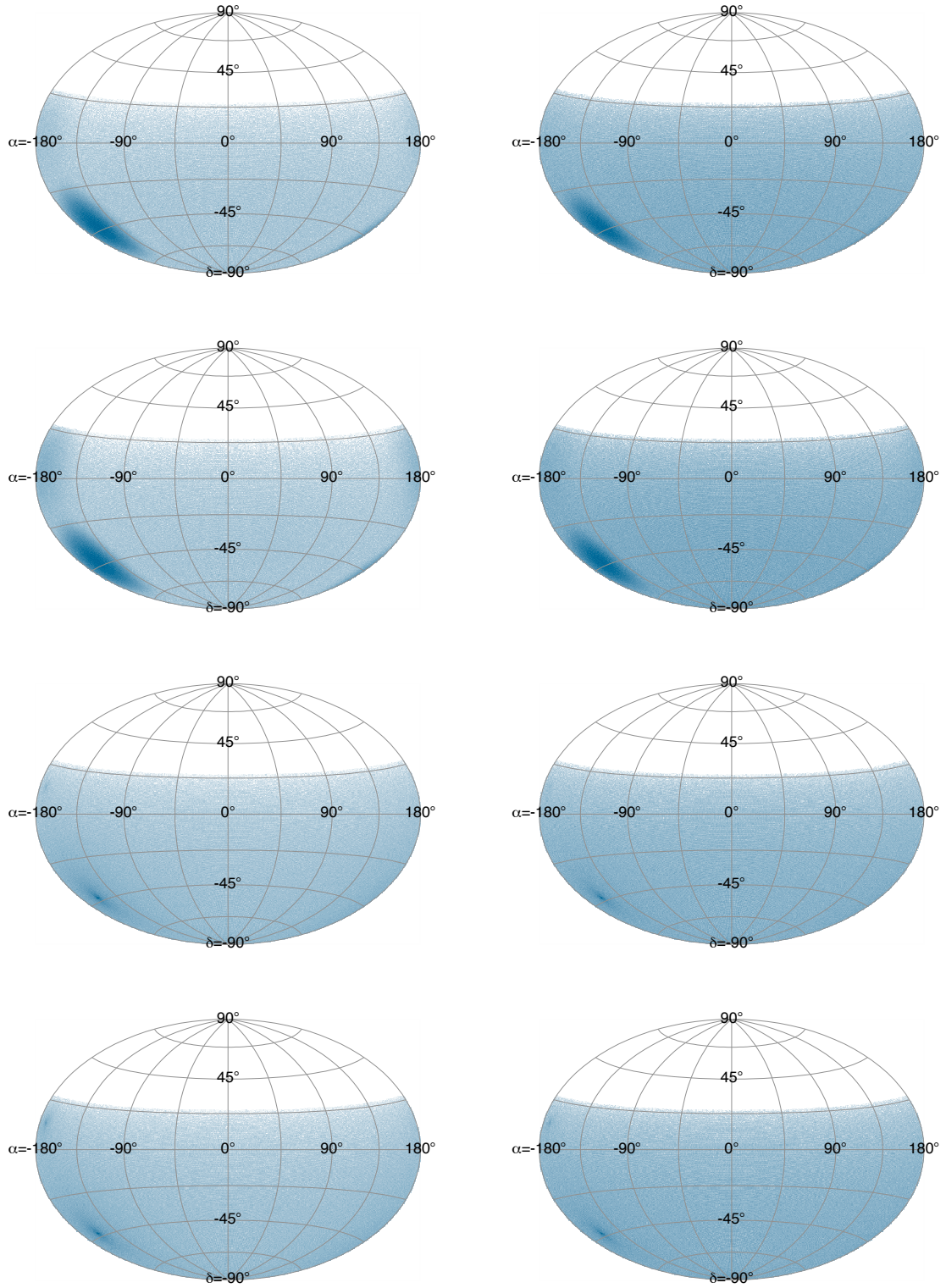


Figure D.3: MC maps based on radio loud galaxies using different parameter sets. On the left the isotropic contribution is 40 % and on the right 80 %. *Top Panel:* The relative strength follows the prediction of model S with a Gaussian smearing. *Upper Middle Panel:* The relative strength follows the prediction of model A with a Gaussian smearing. *Lower Middle Panel:* The relative strength follows the prediction of model S with an inverse smearing. *Bottom Panel:* The relative strength follows the prediction of model A with an inverse smearing.

E. Optimal Radius of the Cluster Algorithm

E.1 Dependence of the Cluster Algorithm on the Radius

In contrast to the other methods presented the Cluster Algorithm has a free parameter, the maximal radius R . Three ways exist to handle it. The first one is to fix it a priori motivated by reasonable models predicting a certain scale of deflections. However, it is still unclear what the sources are and what the corresponding smearing is. Thus, in the second approach a Monte Carlo study can be performed covering many different models where several radii are tested to find out which one gives the best results on average. The third one includes a scan on actual data, i.e. the radius is optimized on a data set followed by a prescription. Further data will then answer the question if this was the optimal choice and an anisotropy is still present at that scale or if it was a fluctuation and nature differs from this.

We have decided to follow the second approach. The several Monte Carlo sets produced in chapter 6 are analysed with six different radii, 5° to 30° in steps of 5° . As a measure of the sensitivity we use the fraction of analysed sets having a probability $P < 0.01$ for the unweighted (CA I) as well as the weighted algorithm (CA II). The radius providing the best results will be presented again in chapter 6 and then compared to the other methods.

Due to the vast amount of source scenarios considered we simplify it by averaging over all set sizes which are between 20 and 100 to reduce the number of parameters by one. However, there are still many parameters left and hence, it is impossible to show a single histogram displaying all results. Therefore, the results are subdivided so that one histogram can contain the results of more than one parameter set. For every other plot only one parameter is changed, e.g. the smearing angle.

All shown histograms have two things in common, the x-axis gives the fraction of isotropy to the total amount of events in one set in % and the y-axis corresponds to the radius used.

Monte Carlo Maps based on catalogues

In Fig. E.1 the results are presented for all catalogue based maps which have been analysed with CA I. Each of the displayed histograms is split in two whereas the left side contains maps having a Gaussian smearing of 3° and the right one a smearing of 6° . The top row shows the results for the VCV weighted with distance (left) and

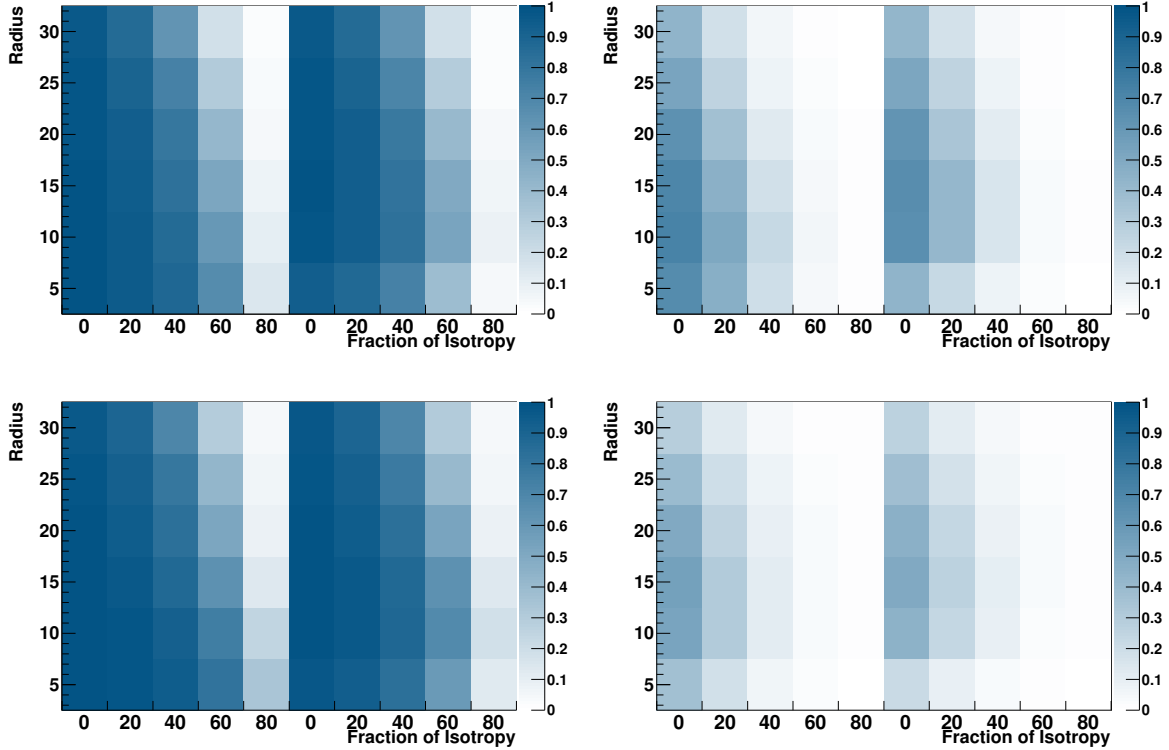


Figure E.1: Displayed is the fraction of sets analysed with CA I giving a probability smaller than 0.01. The top row contains all VCV and the bottom one the corresponding IRAS maps. The left columns represents a distance weighting according to the redshift of the galaxies whereas in the right one each galaxy has the same flux on Earth. On the x-axis of each histogram the contribution of isotropy in % is given. In addition, the histogram is split in 2, the left half corresponds to a smearing of 3° and the right one to 6° .

unweighted (right). Accordingly, the bottom row gives the corresponding results for the maps based on IRAS.

Since we are interested solely in the performance of the method depending on the radius we focus on this parameter. Further discussion regarding the type of the map and the corresponding sensitivity is carried out in chapter 6.

From the displayed maps one can conclude that in the case of a small amount of isotropy and only a few sources the radius is not an important factor. However, when increasing one of these two parameters it can be seen that a radius of $10^\circ - 15^\circ$ is favoured.

In case of the CA II (see Fig. E.2), however, this is not true. Only a minimal radius is required. For all R larger than this value the fraction is nearly constant or only decreasing slightly.

Monte Carlo Maps based on few sources

For this kind of anisotropy two different approaches have been chosen, a model based on radio loud galaxies and one including ellipses at random positions in the field of view of the PAO. The first model has three free parameters (strength, smearing and

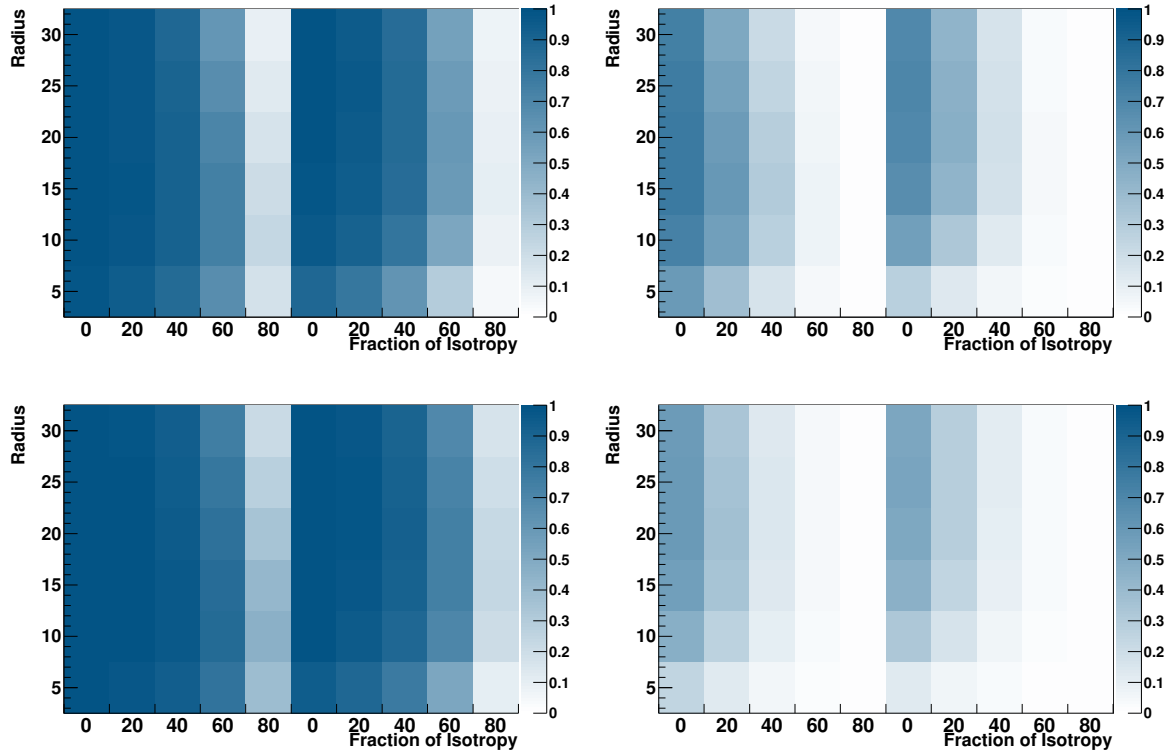


Figure E.2: Displayed is the fraction of sets analysed with CA II giving a probability smaller than 0.01. The top row contains all VCV maps and the bottom one the corresponding IRAS maps. The left column represents a distance weighting according to the redshift of the galaxies whereas in the right one each galaxy has the same flux on Earth. On the x-axis of each histogram the contribution of isotropy in % is given. In addition, the histogram is split in 2, the left half corresponds to a smearing of 3° and the right one to 6° .

isotropic contribution), while the second one has only two (size of the ellipse and isotropic contribution).

The strength corresponds to the column of the histograms in Fig. E.3, left model A, right model S. Within each histogram a similar subdivision is made for the smearing, left (Inverse) and right (Gaussian), which in turn is again subdivided corresponding to the amount of isotropy in the data set. The top row corresponds to CA I and the bottom row to CA II.

It is worth noting, that by changing the smearing function the sensitivity of both variants drop drastically. In contrast to the Gaussian where nearly all sets have small probabilities, just 40 % or less show a similar behaviour. In addition, the influence of the increasing background is stronger leading to an insensitivity for contributions of 60 % or larger. However, this effect is much weaker for the Gaussian smearing.

The sizes of the simulated ellipses are $((10^\circ, 10^\circ), (15^\circ, 5^\circ), (20^\circ, 10^\circ))$ and correspond to three parts in each histogram in Fig. E.4 (the left one corresponds to CA I and the right to CA II) which are divided in the same way corresponding to their isotropic contribution.

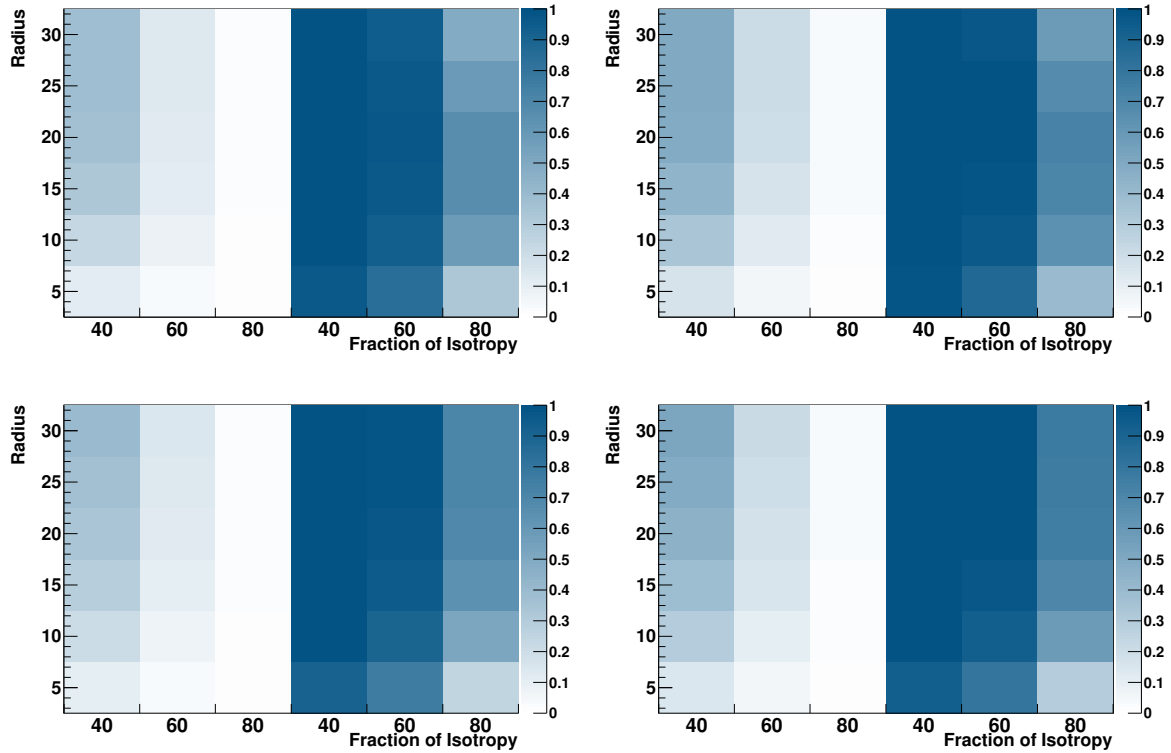


Figure E.3: Fraction of sets based on radio loud galaxies giving a probability smaller than 0.01 depending on radius. Each histogram is divided into two halves, inverse smearing on the left and Gaussian on the right. The strength model A corresponds to the left column of histograms and the right to model S. The isotropic contribution is given on x-axis in %. *Top Panel:* Sets analysed with CA I. *Bottom Panel:* Sets analysed with CA II.

Regarding this model, both variants behave as expected. By increasing the background the sensitivity is weakened whereupon the influence of the size of the ellipses is negligible.

Furthermore, the observed behaviour regarding the dependency on the radius found in the catalogue based maps can be confirmed here for both source models.

Monte Carlo Maps based on multipoles

Due to the design of the methods, one expects that they should be rather insensitive to this kind of anisotropy: particularly at small radii a multipole looks very similar to isotropy. However, by increasing the radius one should be able to see some of the multipoles depending on how strong and located the maxima are.

The results for the analysed multipole sets are displayed in Fig. E.5 (CA I the left histograms and CA II the right one with the type of the dipole on the x-axis). For both methods the results are very similar.

A remarkable sensitivity is reached for the dipole in x direction as well as for the quadrupole in x and z direction, respectively.

With the density maps presented in chapter 6 (cf. Fig. 6.8) this result can be ex-

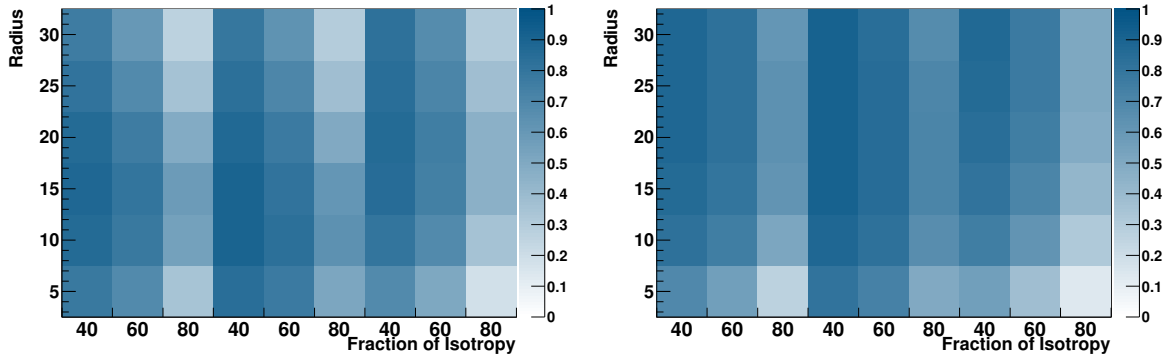


Figure E.4: Fraction of sets based on ellipses giving a probability smaller than 0.01 depending on radius. Each histogram is divided in three parts, the left corresponds to a long axis of 10° , the middle to 15° and the right to 20° . Each of these is again divided corresponding to the contribution of isotropy. *Left Panel:* Sets analysed with CA I. *Right Panel:* Sets analysed with CA II.

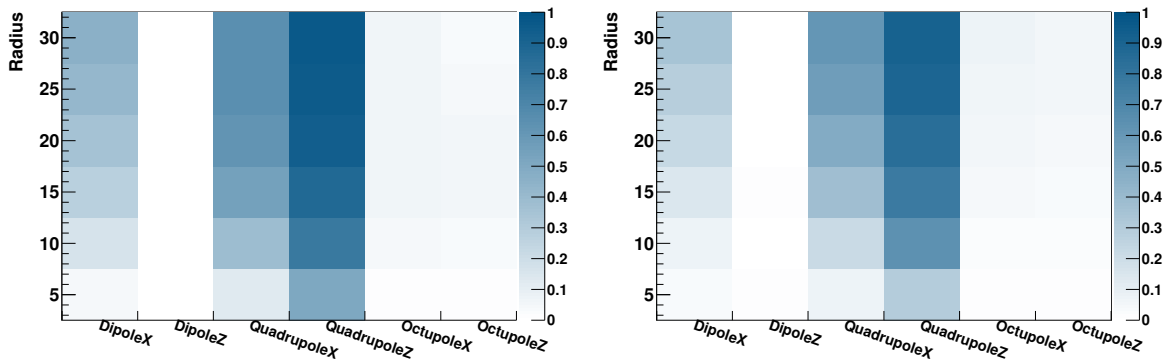


Figure E.5: Fraction of sets based on multipoles giving a probability smaller than 0.01 depending on radius. *Left Panel:* Sets analysed with CA I. *Right Panel:* Sets analysed with CA II.

plained: Only those maps showing areas of clear overabundance of events can give rise to significant contributions to the Cluster Sum. This is particularly visible for the quadrupole in z direction which gives indeed the best results. The others are to close to isotropy in shape and structure and thus, they are not recognized.

E.1.1 Summary

For each radius the amount of sets giving a probability smaller than 0.01 has been calculated and the results are shown in Tab. E.1. Since, CA I as well as CA II are not designed to be sensitive to multipoles (MP) also the result excluding these maps is given.

As a comparison we also give the result for a fixed event set size of 100 in Tab. E.2 to check if an averaging over all set sizes leads to a domination of the results by small sets.

Table E.1: Given are the overall fraction of sets having probabilities below 0.01 for all sizes (20 to 100 in steps of 20). Due to the know insensitivity of both algorithms to multipoles (MP), results with and without including multipoles are presented.

Method	Fraction of sets below $P = 0.01$ for different R					
	5°	10°	15°	20°	25°	30°
CA I (w/ MP)	0.386623	0.481698	0.510503	0.505536	0.479242	0.44943
CA II (w/ MP)	0.337495	0.451382	0.503315	0.526434	0.537056	0.54024
CA I (w/o MP)	0.474711	0.562559	0.579351	0.560455	0.51989	0.47781
CA II (w/o MP)	0.425093	0.546661	0.593718	0.609283	0.612353	0.608199

Table E.2: Given are the overall fraction of sets having probabilities below 0.01 for all a set size of 100. Due to the know insensitivity of both algorithms to multipoles (MP), results with and without including multipoles are presented.

Method	Fraction of sets below $P = 0.01$ for different R					
	5°	10°	15°	20°	25°	30°
CA I (w/ MP)	0.553474	0.653161	0.674742	0.660188	0.639066	0.599708
CA II (w/ MP)	0.501151	0.621098	0.662995	0.681002	0.687854	0.691335
CA I (w/o MP)	0.661437	0.742859	0.751172	0.725079	0.690105	0.638205
CA II (w/o MP)	0.619391	0.732125	0.761087	0.768158	0.767022	0.763969

It is worth noting that for CA I and CA II the results are nearly the same considering all set sizes as well as only the one containing 100 events. The largest value is obtained for a radius $R = 15^\circ$ followed by $R = 10^\circ$ (CA I) and $R = 20^\circ$ followed by $R = 15^\circ$ (CA II) independent of the choice to include or exclude the maps based on multipoles. Since the difference in both cases is only of the order of 2-4 % and due to simplicity we keep only one radius for both methods, $R = 15^\circ$.

However, this result is biased of course by the arbitrary choice of MC maps considered. Note that a different optimal radius would be chosen in the end when used only those based on multipoles. As an alternative we consider a formula 6.5 presented in chapter 6 which gives the smearing scale based on assumptions and/or measurements of distance of the source, the charge of the primary particle, intergalactic B-field and its coherence length. In this context a maximal radius of 10° was emphasized using a strength of the magnetic field to be a factor two larger than recent measurements which leads in turn to a decrease of the smearing by a factor of two. If this is the true average smearing then the optimal radius of CA I would be 5° or 10° whereas for CA II this only sets the minimal appropriate radius. From the MC study we see that afterwards the sensitivity only varies slightly. Hence, we accept the optimal radius found in this MC study and use it to compare CA I as well as CA II to the other methods.

F. Results of the Monte Carlo Study

In chapter 6 many different source models have been simulated and analysed whereas only those results are given being averaged over all set sizes (20, 40, ..., 100). Here, we now list in more detail the results of the MC study. The tables are ordered according to the introduction of the models; starting with the catalogue based scenarios followed by the single source models and ending with the multipoles.

F.1 Maps based on catalogues

From Tab. F.1 and F.2 one can see that in the case of a pure contribution from a few strong galaxies already 20 events are enough to discover the anisotropic distribution of arrival directions of cosmic rays. By increasing the fraction of the isotropic background to values of 40 % or even more the signal is diluted and only 50 % of all sets are still recognized.

Considering the development of the power of the methods for larger event set sizes one can see that for 100 events the ability to find an anisotropic signal has increased as expected. All methods are nearly on the same level, only the MST shows a bit earlier a weakening of its power when adding more isotropically distributed events. The TwoPt correlation functions as well as the cluster algorithms are still properly working in the presence of 80 % background.

However, this result is weakened if the number of sources contributing in the same way, i.e. having the same strength on Earth, is enlarged. Only the CA II is still above 90 % for an isotropic background of 20 % in the case of 100 events. The others are below this value and drop including the CA II to the percent level when the signal is of the order of 20 %.

If we can be sure that the signal should look like anything covered by these models, then the CA II gives the best results averaged over all parameter sets.

Table F.1: Presented are the fractions of sets based on the IRAS with distance weighting giving a probability smaller than 0.01 for a set size of 20.

IRAS with distance weighting										
Smearing	3°					6°				
Isotropy [%]	0	20	40	60	80	0	20	40	60	80
MST	0.90	0.54	0.21	0.07	0.02	0.84	0.46	0.18	0.05	0.02
TwoPt	0.98	0.84	0.54	0.20	0.03	0.95	0.75	0.40	0.13	0.02
TwoPtPlus	0.85	0.56	0.27	0.08	0.02	0.72	0.40	0.17	0.05	0.01
CA I	0.98	0.82	0.50	0.18	0.03	0.98	0.81	0.49	0.17	0.03
CA II	1.00	0.97	0.81	0.47	0.11	0.98	0.88	0.62	0.28	0.05

Table F.2: Presented are the fractions of sets based on the VCV with distance weighting giving a probability smaller than 0.01 for a set size of 20.

VCV with distance weighting										
Smearing	3°					6°				
Isotropy [%]	0	20	40	60	80	0	20	40	60	80
MST	0.90	0.52	0.20	0.05	0.01	0.86	0.46	0.17	0.05	0.01
TwoPt	0.96	0.72	0.36	0.11	0.02	0.91	0.60	0.27	0.07	0.02
TwoPtPlus	0.86	0.51	0.20	0.05	0.01	0.78	0.42	0.15	0.04	0.01
CA I	0.96	0.71	0.37	0.11	0.02	0.95	0.71	0.35	0.10	0.02
CA II	0.99	0.90	0.62	0.27	0.05	0.95	0.75	0.43	0.15	0.03

Table F.3: Presented are the fractions of sets based on the IRAS with distance weighting giving a probability smaller than 0.01 for a set size of 40.

IRAS with distance weighting										
Smearing	3°					6°				
Isotropy [%]	0	20	40	60	80	0	20	40	60	80
MST	1.00	0.93	0.60	0.19	0.03	1.00	0.89	0.49	0.13	0.03
TwoPt	1.00	0.99	0.84	0.39	0.05	1.00	0.98	0.75	0.27	0.03
TwoPtPlus	1.00	0.97	0.72	0.27	0.04	1.00	0.91	0.55	0.17	0.03
CA I	1.00	1.00	0.89	0.46	0.07	1.00	1.00	0.89	0.46	0.07
CA II	1.00	1.00	0.99	0.84	0.26	1.00	1.00	0.95	0.64	0.13

Table F.4: Presented are the fractions of sets based on the VCV with distance weighting giving a probability smaller than 0.01 for a set size of 40.

VCV with distance weighting										
Smearing	3°					6°				
Isotropy [%]	0	20	40	60	80	0	20	40	60	80
MST	1.00	0.91	0.53	0.14	0.02	1.00	0.86	0.43	0.11	0.02
TwoPt	1.00	0.98	0.72	0.24	0.03	1.00	0.95	0.60	0.16	0.02
TwoPtPlus	1.00	0.98	0.73	0.25	0.03	1.00	0.96	0.64	0.19	0.03
CA I	1.00	0.99	0.80	0.30	0.04	1.00	0.99	0.77	0.27	0.04
CA II	1.00	1.00	0.97	0.62	0.11	1.00	0.99	0.86	0.40	0.06

Table F.5: Presented are the fractions of sets based on the IRAS with distance weighting giving a probability smaller than 0.01 for a set size of 60.

IRAS with distance weighting										
Smearing	3°					6°				
Isotropy [%]	0	20	40	60	80	0	20	40	60	80
MST	1.00	1.00	0.83	0.35	0.05	1.00	0.99	0.74	0.27	0.04
TwoPt	1.00	1.00	0.94	0.51	0.06	1.00	1.00	0.90	0.41	0.04
TwoPtPlus	1.00	1.00	0.93	0.50	0.07	1.00	0.99	0.84	0.34	0.05
CA I	1.00	1.00	0.99	0.73	0.14	1.00	1.00	0.99	0.73	0.13
CA II	1.00	1.00	1.00	0.96	0.43	1.00	1.00	1.00	0.86	0.23

Table F.6: Presented are the fractions of sets based on the VCV with distance weighting giving a probability smaller than 0.01 for a set size of 60.

VCV with distance weighting										
Smearing	3°					6°				
Isotropy [%]	0	20	40	60	80	0	20	40	60	80
MST	1.00	0.99	0.79	0.28	0.04	1.00	0.98	0.67	0.20	0.03
TwoPt	1.00	1.00	0.88	0.37	0.04	1.00	0.99	0.81	0.27	0.03
TwoPtPlus	1.00	1.00	0.93	0.45	0.05	1.00	1.00	0.88	0.38	0.04
CA I	1.00	1.00	0.97	0.57	0.07	1.00	1.00	0.96	0.54	0.06
CA II	1.00	1.00	1.00	0.86	0.19	1.00	1.00	0.98	0.65	0.09

Table F.7: Presented are the fractions of sets based on the IRAS with distance weighting giving a probability smaller than 0.01 for a set size of 80.

IRAS with distance weighting										
Smearing	3°					6°				
Isotropy [%]	0	20	40	60	80	0	20	40	60	80
MST	1.00	1.00	0.94	0.50	0.08	1.00	1.00	0.87	0.37	0.05
TwoPt	1.00	1.00	0.98	0.57	0.06	1.00	1.00	0.94	0.45	0.04
TwoPtPlus	1.00	1.00	0.98	0.66	0.10	1.00	1.00	0.94	0.49	0.07
CA I	1.00	1.00	1.00	0.89	0.20	1.00	1.00	1.00	0.88	0.20
CA II	1.00	1.00	1.00	0.99	0.59	1.00	1.00	1.00	0.96	0.35

Table F.8: Presented are the fractions of sets based on the VCV with distance weighting giving a probability smaller than 0.01 for a set size of 80.

VCV with distance weighting										
Smearing	3°					6°				
Isotropy [%]	0	20	40	60	80	0	20	40	60	80
MST	1.00	1.00	0.92	0.42	0.05	1.00	1.00	0.83	0.29	0.04
TwoPt	1.00	1.00	0.94	0.41	0.04	1.00	1.00	0.89	0.33	0.03
TwoPtPlus	1.00	1.00	0.98	0.62	0.08	1.00	1.00	0.97	0.54	0.06
CA I	1.00	1.00	1.00	0.75	0.11	1.00	1.00	0.99	0.73	0.10
CA II	1.00	1.00	1.00	0.96	0.30	1.00	1.00	1.00	0.84	0.16

Table F.9: Presented are the fractions of sets based on the IRAS with distance weighting giving a probability smaller than 0.01 for a set size of 100.

IRAS with distance weighting										
Smearing	3°					6°				
Isotropy [%]	0	20	40	60	80	0	20	40	60	80
MST	1.00	1.00	0.99	0.63	0.10	1.00	1.00	0.94	0.48	0.06
TwoPt	1.00	1.00	0.99	0.67	0.07	1.00	1.00	0.98	0.56	0.04
TwoPtPlus	1.00	1.00	1.00	0.82	0.13	1.00	1.00	0.99	0.66	0.08
CA I	1.00	1.00	1.00	0.97	0.26	1.00	1.00	1.00	0.95	0.25
CA II	1.00	1.00	1.00	1.00	0.69	1.00	1.00	1.00	0.99	0.44

Table F.10: Presented are the fractions of sets based on the VCV with distance weighting giving a probability smaller than 0.01 for a set size of 100.

VCV with distance weighting										
Smearing	3°					6°				
Isotropy [%]	0	20	40	60	80	0	20	40	60	80
MST	1.00	1.00	0.98	0.56	0.07	1.00	1.00	0.92	0.39	0.04
TwoPt	1.00	1.00	0.98	0.53	0.05	1.00	1.00	0.96	0.44	0.04
TwoPtPlus	1.00	1.00	1.00	0.80	0.11	1.00	1.00	1.00	0.73	0.09
CA I	1.00	1.00	1.00	0.86	0.12	1.00	1.00	1.00	0.84	0.12
CA II	1.00	1.00	1.00	0.99	0.37	1.00	1.00	1.00	0.93	0.20

Table F.11: Presented are the fractions of sets based on the IRAS with a flat weighting giving a probability smaller than 0.01 for a set size of 20.

IRAS with flat weighting										
Smearing	3°					6°				
Isotropy [%]	0	20	40	60	80	0	20	40	60	80
MST	0.05	0.03	0.02	0.01	0.01	0.05	0.03	0.02	0.01	0.01
TwoPt	0.04	0.03	0.01	0.01	0.01	0.03	0.02	0.02	0.01	0.01
TwoPtPlus	0.02	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
CA I	0.09	0.05	0.03	0.01	0.01	0.08	0.04	0.02	0.01	0.01
CA II	0.10	0.06	0.03	0.02	0.01	0.08	0.04	0.03	0.02	0.01

Table F.12: Presented are the fractions of sets based on the VCV with a flat weighting giving a probability smaller than 0.01 for a set size of 20.

VCV with flat weighting										
Smearing	3°					6°				
Isotropy [%]	0	20	40	60	80	0	20	40	60	80
MST	0.10	0.04	0.02	0.01	0.01	0.09	0.04	0.02	0.01	0.01
TwoPt	0.24	0.13	0.06	0.02	0.01	0.21	0.11	0.05	0.02	0.01
TwoPtPlus	0.05	0.02	0.01	0.01	0.01	0.04	0.02	0.01	0.01	0.01
CA I	0.17	0.08	0.04	0.01	0.01	0.14	0.07	0.03	0.01	0.01
CA II	0.23	0.12	0.06	0.02	0.01	0.14	0.08	0.04	0.02	0.01

Table F.13: Presented are the fractions of sets based on the IRAS with a flat weighting giving a probability smaller than 0.01 for a set size of 40.

IRAS with flat weighting										
Smearing	3°					6°				
Isotropy [%]	0	20	40	60	80	0	20	40	60	80
MST	0.24	0.10	0.04	0.02	0.01	0.20	0.08	0.04	0.02	0.01
TwoPt	0.08	0.03	0.02	0.01	0.01	0.07	0.03	0.02	0.01	0.01
TwoPtPlus	0.06	0.03	0.02	0.01	0.01	0.05	0.03	0.02	0.01	0.01
CA I	0.30	0.14	0.05	0.02	0.01	0.26	0.12	0.05	0.02	0.01
CA II	0.32	0.16	0.07	0.03	0.01	0.22	0.11	0.05	0.03	0.01

Table F.14: Presented are the fractions of sets based on the VCV with a flat weighting giving a probability smaller than 0.01 for a set size of 40.

VCV with flat weighting										
Smearing	3°					6°				
Isotropy [%]	0	20	40	60	80	0	20	40	60	80
MST	0.47	0.17	0.06	0.02	0.01	0.39	0.14	0.04	0.02	0.01
TwoPt	0.57	0.29	0.11	0.03	0.01	0.51	0.25	0.09	0.03	0.01
TwoPtPlus	0.30	0.13	0.05	0.02	0.01	0.26	0.12	0.04	0.01	0.01
CA I	0.52	0.23	0.08	0.03	0.01	0.46	0.20	0.07	0.01	0.01
CA II	0.68	0.37	0.15	0.04	0.01	0.46	0.22	0.08	0.03	0.01

Table F.15: Presented are the fractions of sets based on the IRAS with a flat weighting giving a probability smaller than 0.01 for a set size of 60.

IRAS with flat weighting										
Smearing	3°					6°				
Isotropy [%]	0	20	40	60	80	0	20	40	60	80
MST	0.51	0.21	0.08	0.03	0.01	0.42	0.17	0.06	0.02	0.01
TwoPt	0.15	0.05	0.02	0.01	0.01	0.11	0.04	0.01	0.01	0.01
TwoPtPlus	0.13	0.07	0.04	0.02	0.01	0.11	0.06	0.03	0.02	0.01
CA I	0.60	0.29	0.11	0.04	0.02	0.54	0.26	0.10	0.04	0.01
CA II	0.60	0.30	0.12	0.04	0.02	0.45	0.22	0.08	0.03	0.01

Table F.16: Presented are the fractions of sets based on the VCV with a flat weighting giving a probability smaller than 0.01 for a set size of 60.

VCV with flat weighting										
Smearing	3°					6°				
Isotropy [%]	0	20	40	60	80	0	20	40	60	80
MST	0.83	0.39	0.12	0.03	0.01	0.71	0.29	0.09	0.02	0.01
TwoPt	0.81	0.46	0.17	0.04	0.01	0.75	0.41	0.15	0.04	0.02
TwoPtPlus	0.57	0.28	0.10	0.03	0.01	0.50	0.24	0.09	0.03	0.01
CA I	0.85	0.48	0.17	0.04	0.01	0.79	0.42	0.15	0.04	0.01
CA II	0.94	0.66	0.28	0.07	0.01	0.78	0.43	0.16	0.04	0.01

Table F.17: Presented are the fractions of sets based on the IRAS with a flat weighting giving a probability smaller than 0.01 for a set size of 80.

IRAS with flat weighting										
Smearing	3°					6°				
Isotropy [%]	0	20	40	60	80	0	20	40	60	80
MST	0.77	0.34	0.11	0.03	0.01	0.64	0.26	0.09	0.03	0.01
TwoPt	0.18	0.06	0.02	0.01	0.01	0.13	0.05	0.012	0.01	0.01
TwoPtPlus	0.22	0.10	0.05	0.02	0.01	0.19	0.09	0.04	0.02	0.01
CA I	0.83	0.47	0.20	0.05	0.01	0.75	0.40	0.15	0.05	0.02
CA II	0.85	0.53	0.21	0.06	0.02	0.70	0.36	0.14	0.04	0.02

Table F.18: Presented are the fractions of sets based on the VCV with a flat weighting giving a probability smaller than 0.01 for a set size of 80.

VCV with flat weighting										
Smearing	3°					6°				
Isotropy [%]	0	20	40	60	80	0	20	40	60	80
MST	0.98	0.62	0.21	0.05	0.01	0.91	0.47	0.14	0.03	0.01
TwoPt	0.92	0.60	0.23	0.05	0.01	0.88	0.54	0.19	0.05	0.01
TwoPtPlus	0.78	0.44	0.17	0.05	0.02	0.72	0.39	0.14	0.04	0.02
CA I	0.97	0.69	0.28	0.06	0.01	0.94	0.62	0.24	0.05	0.01
CA II	1.00	0.88	0.45	0.11	0.02	0.95	0.66	0.27	0.06	0.01

Table F.19: Presented are the fractions of sets based on the IRAS with a flat weighting giving a probability smaller than 0.01 for a set size of 100.

IRAS with flat weighting										
Smearing	3°					6°				
Isotropy [%]	0	20	40	60	80	0	20	40	60	80
MST	0.92	0.47	0.14	0.04	0.01	0.80	0.35	0.12	0.04	0.01
TwoPt	0.26	0.09	0.03	0.01	0.01	0.19	0.07	0.03	0.02	0.01
TwoPtPlus	0.35	0.17	0.05	0.02	0.01	0.32	0.14	0.05	0.02	0.01
CA I	0.92	0.59	0.22	0.06	0.01	0.87	0.51	0.19	0.05	0.02
CA II	0.95	0.67	0.28	0.07	0.02	0.83	0.49	0.19	0.05	0.01

Table F.20: Presented are the fractions of sets based on the VCV with a flat weighting giving a probability smaller than 0.01 for a set size of 100.

VCV with flat weighting										
Smearing	3°					6°				
Isotropy [%]	0	20	40	60	80	0	20	40	60	80
MST	1.00	0.79	0.31	0.07	0.01	0.98	0.62	0.21	0.04	0.01
TwoPt	0.97	0.76	0.33	0.08	0.02	0.96	0.69	0.27	0.07	0.02
TwoPtPlus	0.94	0.65	0.27	0.06	0.02	0.90	0.59	0.23	0.05	0.01
CA I	0.99	0.82	0.36	0.07	0.01	0.98	0.76	0.30	0.05	0.01
CA II	1.00	0.96	0.60	0.15	0.02	0.99	0.81	0.36	0.08	0.02

F.2 Maps based on single sources

Considering this type of maps one assumes that by design the CA should be the most sensitive method. Since, the shape of the event distribution and the number of sources matches perfectly the incorporated assumptions. The corresponding tables are Tab. F.21-F.35.

Radio loud galaxies

Although, this is true for the model using a Gaussian smearing for the radio loud galaxies, a change of the distribution function towards the Inverse smearing leads to a drastic drop in sensitivity which only increases up to values of 50 % with the smallest background for strength model A. In the case of model S where the fraction coming from Cen A is further enhanced towards roughly 70 %.

The other methods, especially the TwoPtPlus, show a comparable behaviour only slightly worse in sensitivity. Consequently, a single source whose cosmic rays are smeared with a Gaussian can definitely be identified for a small number of events. And for an even larger background this can be achieved at 100 events.

Furthermore, a change in the deflection function leading to a distribution around a source following the Inverse model will cause a loss in performance for all methods.

Ellipse model

However, a flat distribution around a source has no significant effect on the sensitivity of the CA. At a size of 100 events all simulated event sets are identified when compared to the isotropic expectations in case of a low background. When increasing it only the CA is as good as before while MST, TwoPt and TwoPtPlus show worse behaviour which is expected due to the strong influence of isotropically distributed events.

F.3 Maps based on multipoles

The last type of MC maps produced are those based on multipoles (see Tab. F.36-F.40). From the shape of these one can argue that the TwoPt and MST methods should be more sensitive than the cluster algorithms. They take into account the complete available sky whereas the CA focuses on angular scales well below the size of considered multipoles.

Table F.21: Presented are the fractions of sets based on the radio loud galaxies with a Gaussian smearing giving a probability smaller than 0.01 for a set size of 20.

Radio loud galaxies with a Gaussian smearing						
strength model	A			S		
Isotropy [%]	40	60	80	40	60	80
MST	0.70	0.24	0.05	0.81	0.31	0.05
TwoPt	0.99	0.52	0.03	1.00	0.65	0.04
TwoPtPlus	0.89	0.25	0.02	0.96	0.32	0.02
CA I	1.00	0.85	0.10	1.00	0.94	0.12
CA II	0.99	0.79	0.11	1.00	0.88	0.13

Table F.22: Presented are the fractions of sets based on the radio loud galaxies with an Inverse smearing giving a probability smaller than 0.01 for a set size of 20.

Radio loud galaxies with an Inverse smearing						
strength model	A			S		
Isotropy [%]	40	60	80	40	60	80
MST	0.05	0.03	0.01	0.07	0.03	0.01
TwoPt	0.06	0.02	0.01	0.07	0.03	0.01
TwoPtPlus	0.02	0.01	0.01	0.03	0.01	0.01
CA I	0.06	0.03	0.01	0.09	0.03	0.01
CA II	0.06	0.03	0.01	0.08	0.03	0.01

Table F.23: Presented are the fractions of sets based on the radio loud galaxies with a Gaussian smearing giving a probability smaller than 0.01 for a set size of 40.

Radio loud galaxies with a Gaussian smearing						
strength model	A			S		
Isotropy [%]	40	60	80	40	60	80
MST	0.95	0.46	0.07	0.99	0.58	0.10
TwoPt	1.00	0.94	0.11	1	0.99	0.14
TwoPtPlus	1.00	0.93	0.10	1	0.96	0.11
CA I	1.00	1.00	0.45	1.00	1.00	0.56
CA II	1.00	1.00	0.44	1.00	1.00	0.54

However, as can be seen from the corresponding tables the quadrupole in z direction forms an exception to these assumptions. It shows an enhanced sensitivity for all methods. This can be explained by its shape which includes several well separated maxima which can be found e.g. by the CA especially for larger radii. In addition, for the TwoPt functions further types like the dipole in z direction can be clearly identified for 100 events in a set.

Table F.24: Presented are the fractions of sets based on the radio loud galaxies with an Inverse smearing giving a probability smaller than 0.01 for a set size of 40.

Radio loud galaxies with an Inverse smearing						
strength model	A			S		
Isotropy [%]	40	60	80	40	60	80
MST	0.07	0.03	0.01	0.10	0.04	0.02
TwoPt	0.13	0.04	0.01	0.19	0.06	0.02
TwoPtPlus	0.12	0.03	0.01	0.14	0.04	0.01
CA I	0.16	0.05	0.02	0.24	0.08	0.02
CA II	0.15	0.05	0.02	0.22	0.08	0.02

Table F.25: Presented are the fractions of sets based on the radio loud galaxies with a Gaussian smearing giving a probability smaller than 0.01 for a set size of 60.

Radio loud galaxies with a Gaussian smearing						
strength model	A			S		
Isotropy [%]	40	60	80	40	60	80
MST	0.99	0.67	0.11	1.00	0.78	0.14
TwoPt	1.00	0.99	0.19	1.00	1.00	0.26
TwoPtPlus	1.00	1.00	0.24	1	1.00	0.28
CA I	1.00	1.00	0.81	1.00	1.00	0.91
CA II	1.00	1.00	0.75	1.00	1.00	0.86

Table F.26: Presented are the fractions of sets based on the radio loud galaxies with an Inverse smearing giving a probability smaller than 0.01 for a set size of 60.

Radio loud galaxies with an Inverse smearing						
strength model	A			S		
Isotropy [%]	40	60	80	40	60	80
MST	0.09	0.04	0.02	0.13	0.05	0.02
TwoPt	0.21	0.06	0.02	0.33	0.10	0.02
TwoPtPlus	0.23	0.06	0.02	0.28	0.08	0.02
CA I	0.32	0.11	0.02	0.46	0.16	0.03
CA II	0.27	0.09	0.02	0.39	0.14	0.03

Table F.27: Presented are the fractions of sets based on the radio loud galaxies with a Gaussian smearing giving a probability smaller than 0.01 for a set size of 80.

Radio loud galaxies with a Gaussian smearing						
strength model	A			S		
Isotropy [%]	40	60	80	40	60	80
MST	1.00	0.81	0.14	1	0.89	0.19
TwoPt	1.00	1.00	0.26	1.00	1.00	0.36
TwoPtPlus	1.00	1.00	0.41	1.00	1.00	0.48
CA I	1.00	1.00	0.95	1.00	1.00	0.99
CA II	1.00	1.00	0.93	1.00	1.00	0.98

Table F.28: Presented are the fractions of sets based on the radio loud galaxies with an Inverse smearing giving a probability smaller than 0.01 for a set size of 80.

Radio loud galaxies with an Inverse smearing						
strength model	A			S		
Isotropy [%]	40	60	80	40	60	80
MST	0.11	0.04	0.02	0.17	0.06	0.02
TwoPt	0.29	0.09	0.02	0.45	0.13	0.03
TwoPtPlus	0.41	0.12	0.02	0.46	0.14	0.03
CA I	0.48	0.17	0.03	0.65	0.25	0.04
CA II	0.42	0.15	0.03	0.57	0.22	0.04

Table F.29: Presented are the fractions of sets based on the radio loud galaxies with a Gaussian smearing giving a probability smaller than 0.01 for a set size of 100.

Radio loud galaxies with a Gaussian smearing						
strength model	A			S		
Isotropy [%]	40	60	80	40	60	80
MST	1.00	0.88	0.18	1.00	0.95	0.24
TwoPt	1.00	1.00	0.35	1.00	1.00	0.47
TwoPtPlus	1.00	1.00	0.61	1.00	1.00	0.68
CA I	1.00	1.00	0.99	1.00	1.00	1.00
CA II	1.00	1.00	0.98	1.00	1.00	1.00

Table F.30: Presented are the fractions of sets based on the radio loud galaxies with an Inverse smearing giving a probability smaller than 0.01 for a set size of 100.

Radio loud galaxies with an Inverse smearing						
strength model	A			S		
Isotropy [%]	40	60	80	40	60	80
MST	0.13	0.05	0.02	0.20	0.08	0.02
TwoPt	0.39	0.11	0.02	0.56	0.18	0.03
TwoPtPlus	0.55	0.16	0.03	0.61	0.20	0.03
CA I	0.60	0.22	0.04	0.77	0.33	0.05
CA II	0.53	0.19	0.04	0.70	0.29	0.04

Table F.31: Presented are the fractions of sets based on the Ellipse giving a probability smaller than 0.01 for a set size of 20.

Ellipses with dimensions									
Axes	10°, 10°			15°, 5°			20°, 10°		
Isotropy [%]	40	60	80	40	60	80	40	60	80
MST	0.35	0.14	0.04	0.44	0.18	0.04	0.38	0.14	0.04
TwoPt	0.15	0.06	0.03	0.24	0.10	0.03	0.23	0.09	0.02
TwoPtPlus	0.09	0.04	0.01	0.13	0.05	0.02	0.09	0.03	0.01
CA I	0.44	0.18	0.05	0.50	0.21	0.05	0.33	0.13	0.04
CA II	0.34	0.16	0.05	0.55	0.27	0.08	0.20	0.10	0.04

Table F.32: Presented are the fractions of sets based on the Ellipse giving a probability smaller than 0.01 for a set size of 40.

Ellipses with dimensions									
Axes	10°, 10°			15°, 5°			20°, 10°		
Isotropy [%]	40	60	80	40	60	80	40	60	80
MST	0.89	0.42	0.09	0.94	0.51	0.10	0.81	0.35	0.07
TwoPt	0.74	0.35	0.07	0.88	0.49	0.10	0.74	0.35	0.07
TwoPtPlus	0.53	0.26	0.06	0.62	0.31	0.07	0.42	0.18	0.05
CA I	1.00	0.84	0.28	1.00	0.87	0.30	0.94	0.64	0.18
CA II	1.00	0.87	0.34	1.00	0.99	0.57	0.87	0.56	0.17

Table F.33: Presented are the fractions of sets based on the Ellipse giving a probability smaller than 0.01 for a set size of 60.

Ellipses with dimensions									
Axes	10°, 10°			15°, 5°			20°, 10°		
Isotropy [%]	40	60	80	40	60	80	40	60	80
MST	0.99	0.70	0.15	1.00	0.78	0.18	0.97	0.57	0.13
TwoPt	0.96	0.67	0.15	0.98	0.79	0.21	0.95	0.63	0.14
TwoPtPlus	0.84	0.58	0.17	0.91	0.68	0.20	0.74	0.43	0.12
CA I	1.00	1.00	0.68	1.00	1.00	0.73	1.00	0.95	0.47
CA II	1.00	1.00	0.74	1.00	1.00	0.93	1.00	0.91	0.41

Table F.34: Presented are the fractions of sets based on the Ellipse giving a probability smaller than 0.01 for a set size of 80.

Ellipses with dimensions									
Axes	10°, 10°			15°, 5°			20°, 10°		
Isotropy [%]	40	60	80	40	60	80	40	60	80
MST	1.00	0.85	0.22	1	0.92	0.26	0.99	0.73	0.16
TwoPt	0.99	0.81	0.24	1.00	0.88	0.30	0.99	0.79	0.22
TwoPtPlus	0.95	0.79	0.33	0.98	0.86	0.38	0.88	0.64	0.24
CA I	1.00	1.00	0.93	1.00	1.00	0.95	1.00	1.00	0.73
CA II	1.00	1.00	0.96	1.00	1.00	1.00	1.00	0.99	0.68

Table F.35: Presented are the fractions of sets based on the Ellipse giving a probability smaller than 0.01 for a set size of 100.

Ellipses with dimensions									
Axes	10°, 10°			15°, 5°			20°, 10°		
Isotropy [%]	40	60	80	40	60	80	40	60	80
MST	1.00	0.92	0.27	1.00	0.96	0.33	1.00	0.81	0.21
TwoPt	0.99	0.87	0.33	1.00	0.92	0.38	1.00	0.87	0.30
TwoPtPlus	0.98	0.89	0.48	1.00	0.94	0.56	0.95	0.77	0.35
CA I	1.00	1.00	0.99	1.00	1.00	0.99	1.00	1.00	0.88
CA II	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.84

Table F.36: Presented are the fractions of sets based on multipoles giving a probability smaller than 0.01 for a set size of 20.

Multipole	Dipole		Quadrupole		Octupole	
	X	Z	X	Z	X	Z
MST	0.10	0.01	0.05	0.83	0.02	0.02
TwoPt	0.09	0.38	0.63	0.93	0.01	0.01
TwoPtPlus	0.07	0.06	0.24	0.15	0.01	0.05
CA I	0.05	0.01	0.10	0.44	0.02	0.02
CA II	0.03	0.01	0.06	0.21	0.02	0.01

Table F.37: Presented are the fractions of sets based on multipoles giving a probability smaller than 0.01 for a set size of 40.

Multipole	Dipole		Quadrupole		Octupole	
	X	Z	X	Z	X	Z
MST	0.18	0.01	0.26	0.97	0.03	0.05
TwoPt	0.27	0.88	0.96	1.00	0.02	0.02
TwoPtPlus	0.49	0.14	0.88	0.43	0.02	0.16
CA I	0.12	0.01	0.29	0.95	0.03	0.03
CA II	0.07	0.01	0.16	0.71	0.03	0.03

Table F.38: Presented are the fractions of sets based on multipoles giving a probability smaller than 0.01 for a set size of 60.

Multipole	Dipole		Quadrupole		Octupole	
	X	Z	X	Z	X	Z
MST	0.28	0.01	0.53	0.99	0.05	0.08
TwoPt	0.47	0.99	1.00	1	0.02	0.05
TwoPtPlus	0.72	0.36	1.00	0.81	0.04	0.43
CA I	0.25	0.01	0.61	1.00	0.06	0.05
CA II	0.12	0.01	0.34	0.96	0.04	0.03

Table F.39: Presented are the fractions of sets based on multipoles giving a probability smaller than 0.01 for a set size of 80.

Multipole	Dipole		Quadrupole		Octupole	
	X	Z	X	Z	X	Z
MST	0.36	0.01	0.72	1.00	0.07	0.12
TwoPt	0.63	1.00	1.00	1.00	0.03	0.07
TwoPtPlus	0.94	0.63	1.00	0.96	0.06	0.69
CA I	0.41	0.01	0.83	1.00	0.09	0.08
CA II	0.22	0.01	0.58	1.00	0.06	0.05

Table F.40: Presented are the fractions of sets based on multipoles giving a probability smaller than 0.01 for a set size of 100.

Multipole	Dipole		Quadrupole		Octupole	
	X	Z	X	Z	X	Z
MST	0.43	0.01	0.85	1.00	0.08	0.14
TwoPt	0.75	1.00	1.00	1.00	0.04	0.13
TwoPtPlus	0.99	0.68	1.00	0.98	0.09	0.77
CA I	0.54	0.01	0.93	1.00	0.11	0.09
CA II	0.32	0.01	0.75	1.00	0.07	0.06

F.4 Summary

In general, all methods show the expected behaviour, i.e. the stronger the signal and the larger the event set size the more sensitive they become. The only exception to this behaviour can be found on maps based on multipole such as the dipole in z direction and the CAs. Here, the fraction of identified sets stays at the level of 0.01. Consequently, these methods are completely insensitive to this kind of anisotropy. However, other multipoles exist, e.g. the quadrupole in z direction which can be distinguished from the Null hypothesis for all sets containing 100 events.

References

- [1] G. DE VAUCOULEURS *et al.*, *Third Reference Catalogue of Bright Galaxies*, 1991.
- [2] O. LAHAV *et al.*, *The supergalactic plane revisited with the Optical Redshift Survey*, *Mon. Not. R. Astron. Soc.*, 312 (2000), pp. 166–176.
- [3] M. P. VERON-CETTY AND P. VERON, *A catalogue of quasars and active nuclei: 12th edition*, *Astron. Astrophys.*, 455 (2006), pp. 773–777.
- [4] V. F. HESS, *Über die beobachtungen der durchdringenden strahlung bei sieben freiballonfahrten*, *Phys. Z.*, 13 (1912).
- [5] C. D. ANDERSON, *The Apparent Existence of Easily Deflectable Positives*, *Science*, 76 (1932), pp. 238–239.
- [6] C. D. ANDERSON AND S. H. NEDDERMEYER, *Cloud Chamber Observations of Cosmic Rays at 4300 Meters Elevation and Near Sea-Level*, *Phys. Rev.*, 50 (1936), pp. 263–271.
- [7] K. WERNER AND T. PIEROG, *Extended Air Shower Simulations Based on EPOS*, *AIP Conf. Proc.*, 928 (2007), pp. 111–117.
- [8] S. OSTAPCHENKO, *QGSJET-II: Towards reliable description of very high energy hadronic interactions*, *Nucl.Phys.Proc.Suppl.*, 151 (2006), pp. 143–146.
- [9] D. HECK *et al.*, *CORSIKA: A Monte Carlo code to simulate extensive air showers*. FZKA-6019.
- [10] T. PIEROG *et al.*, *Latest Results from the Air Shower Simulation Programs CORSIKA and CONEX*, (2007).
- [11] S. J. SCIUTTO, *AIRES: A system for air shower simulations (version 2.2.0)*, (1999).
- [12] J. BLUEMER, R. ENGEL, AND J. R. HOERANDEL, *Cosmic Rays from the Knee to the Highest Energies*, *Prog. Part. Nucl. Phys.*, 63 (2009), pp. 293–338.
- [13] P. E. A. SOKOLSKY, *Final Results from the High Resolution Fly’s Eye (HiRes) Experiment*, (2010).
- [14] J. ABRAHAM *et al.*, *Measurement of the Depth of Maximum of Extensive Air Showers above 10^{18} eV*, *Phys. Rev. Lett.*, 104 (2010), p. 091101.

-
- [15] E. FERMI, *On the Origin of the Cosmic Radiation*, Phys. Rev., 75 (1949), pp. 1169–1174.
- [16] T. STANEV, *High energy cosmic rays*, (2004). Berlin, Germany: Springer (2003) 319 p.
- [17] G. F. KRYMSKII, *A regular mechanism for the acceleration of charged particles on the front of a shock wave*, Akademiia Nauk SSSR Doklady, 234 (1977), pp. 1306–1308.
- [18] R. D. BLANDFORD AND J. P. OSTRICKER, *Particle Acceleration by Astrophysical Shocks*, Astrophys. J., 221 (1978), pp. L29–L32.
- [19] T. ANTONI *et al.*, *KASCADE measurements of energy spectra for elemental groups of cosmic rays: Results and open problems*, Astropart.Phys., 24 (2005), pp. 1–25.
- [20] M. BARTAINA *et al.*, *The all particle energy spectrum of KASCADE-Grande in the energy region 10^{16} - 10^{18} eV by means of the N_{ch} - N_{μ} technique*, in 31th Int. Cosmic Ray Conf., 2009.
- [21] R. ABBASI *et al.*, *First observation of the Greisen-Zatsepin-Kuzmin suppression*, Phys.Rev.Lett., 100 (2008), p. 101101.
- [22] J. ABRAHAM *et al.*, *Measurement of the energy spectrum of cosmic rays above 10^{18} eV using the Pierre Auger Observatory*, Phys. Lett., B685 (2010), pp. 239–246.
- [23] A. M. HILLAS, *The Origin of Ultrahigh-Energy Cosmic Rays*, Ann. Rev. Astron. Astrophys., 22 (1984), pp. 425–444.
- [24] D. ALLARD *et al.*, *Cosmogenic neutrinos from the propagation of ultra high energy nuclei*, JCAP, 0609 (2006), p. 005.
- [25] P. ABREU *et al.*, *Update on the correlation of the highest energy cosmic rays with nearby extragalactic matter*, Astropart. Phys., 34 (2010), pp. 314–326.
- [26] J. ABRAHAM *et al.*, *Correlation of the highest-energy cosmic rays with the positions of nearby active galactic nuclei*, Astropart. Phys., 29 (2008), pp. 188–204.
- [27] R. U. ABBASI *et al.*, *Search for Correlations between HiRes Stereo Events and Active Galactic Nuclei*, Astropart. Phys., 30 (2008), pp. 175–179.
- [28] R. U. ABBASI *et al.*, *Analysis of large-scale anisotropy of ultra-high energy cosmic rays in HiRes data*, (2010).
- [29] P. L. BIERMANN *et al.*, *Active Galactic Nuclei: Sources for ultra high energy cosmic rays?*, Nucl. Phys. Proc. Suppl., 190 (2009), pp. 61–78.

-
- [30] M. HARDCASTLE, C. CHEUNG, I. FEAIN, AND L. STAWARZ, *High-energy Particle Acceleration and Production of Ultra-high-energy Cosmic Rays in the Giant Lobes of Centaurus A*, (2008).
- [31] GOPAL-KRISHNA, P. L. BIERMANN, V. DE SOUZA, AND P. J. WIITA, *Ultra-High Energy Cosmic Rays from Centaurus A: Jet Interaction with Gaseous Shells*, (2010).
- [32] P. MESZAROS, *Gamma-Ray Bursts*, Rept. Prog. Phys., 69 (2006), pp. 2259–2322.
- [33] E. WAXMAN, *High energy cosmic-rays from gamma-ray burst sources: A stronger case*, Astrophys. J., 606 (2004), pp. 988–993.
- [34] A. DE RUJULA, *An introduction to cosmic rays and gamma-ray bursts, and to their simple understanding*. Prepared for 21st Les Rencontres de Physique de la Vallée d’Aoste, La Thuile, Aosta, Italy, 4-10 Mar 2007.
- [35] J. ABRAHAM *et al.*, *Upper limit on the cosmic-ray photon fraction at EeV energies from the Pierre Auger Observatory*, Astropart. Phys., 31 (2009), pp. 399–406.
- [36] J. ABRAHAM *et al.*, *Limit on the diffuse flux of ultra-high energy tau neutrinos with the surface detector of the Pierre Auger Observatory*, Phys. Rev., D79 (2009), p. 102001.
- [37] K. GREISEN, *End to the cosmic ray spectrum?*, Phys.Rev.Lett., 16 (1966), pp. 748–750.
- [38] G. ZATSEPIN AND V. KUZMIN, *Upper limit of the spectrum of cosmic rays*, JETP Lett., 4 (1966), pp. 78–80.
- [39] P. BHATTACHARJEE AND G. SIGL, *Extreme energy cosmic rays: Hints to new physics beyond the standard model?*, (2000), pp. 275–299.
- [40] N. HAYASHIDA *et al.*, *Observation of a very energetic cosmic ray well beyond the predicted 2.7-K cutoff in the primary energy spectrum*, Phys. Rev. Lett., 73 (1994), pp. 3491–3494.
- [41] C. WILEMAN, *The Spread in the Arrival Times of Particles in Air-Showers for Photon and Anisotropy Searches above 10 EeV*, PhD thesis, University of Leeds, February 2008.
- [42] T. KAHNIASHVILI *et al.*, *Primordial magnetic field limits from cosmological data*, Phys. Rev., D82 (2010), p. 083005.
- [43] M. ERDMANN AND P. SCHIFFER, *A Method of Measuring Cosmic Magnetic Fields with Ultra High Energy Cosmic Ray Data*, Astropart.Phys., 33 (2010), pp. 201–205.

- [44] E. WAXMAN AND J. MIRALDA-ESCUDE, *Images of bursting sources of high-energy cosmic rays. 1. Effects of magnetic fields*, *Astrophys.J.*, 472 (1996), pp. L89–L92.
- [45] E. SANTOS, *A search for possible anisotropies of cosmic rays with $0.1 \leq E \leq 10$ EeV in the region of the Galactic Centre*, (2007).
- [46] A. H. COMPTON AND I. A. GETTING, *An apparent effect of galactic rotation on the intensity of cosmic rays*, *Phys. Rev.*, 47 (1935), pp. 817–821.
- [47] L. GLEESON AND W. AXFORD, *The compton-getting effect*, *Astrophysics and Space Science*, 2 (1968), pp. 431–437.
- [48] P. ABREU *et al.*, *Search for first harmonic modulation in the right ascension distribution of cosmic rays detected at the Pierre Auger Observatory*, *Astropart. Phys.*, 34 (2011), pp. 627–639.
- [49] N. HAYASHIDA *et al.*, *The Anisotropy of cosmic ray arrival directions around $10^{*}18$ -eV*, *Astropart.Phys.*, 10 (1999), pp. 303–311.
- [50] M. TAKEDA *et al.*, *Small-scale anisotropy of cosmic rays above $10^{*}19$ -eV observed with the Akeno Giant Air Shower Array*, *Astrophys. J.*, 522 (1999), pp. 225–237.
- [51] J. R. T. DE MELLO NETO FOR THE PIERRE AUGER COLLABORATION, *Search for intrinsic anisotropy in the UHECRs data from the Pierre Auger Observatory*, in 31th Int. Cosmic Ray Conf., 2009.
- [52] J. MATTHEWS, *A Heitler model of extensive air showers*, *Astropart. Phys.*, 22 (2005), pp. 387–397.
- [53] M. NAGANO AND A. A. WATSON, *Observations and implications of the ultrahigh-energy cosmic rays*, *Rev. Mod. Phys.*, 72 (2000), pp. 689–732.
- [54] J. A. ABRAHAM *et al.*, *The Fluorescence Detector of the Pierre Auger Observatory*, *Nucl. Instrum. Meth.*, A620 (2010), pp. 227–251.
- [55] M. BOHACOVA *et al.*, *A novel method for the absolute fluorescence yield measurement by AIRFLY*, *Nucl. Instrum. Meth.*, A597 (2008), pp. 55–60.
- [56] C. BONIFAZI AND F. T. P. A. COLLABORATION, *The angular resolution of the Pierre Auger Observatory*, *Nucl.Phys.Proc.Suppl.*, 190 (2009), pp. 20–25.
- [57] B. R. DAWSON, *Hybrid Performance of the Pierre Auger Observatory*, (2007).
- [58] D. KUEMPEL, K.-H. KAMPERT, AND M. RISSE, *Geometry reconstruction of fluorescence detectors revisited*, *Astropart.Phys.*, 30 (2008), pp. 167–174.
- [59] J. ABRAHAM *et al.*, *Trigger and aperture of the surface detector array of the Pierre Auger Observatory*, *Nucl. Instrum. Meth.*, A613 (2010), pp. 29–39.

-
- [60] D. NEWTON, J. KNAPP, AND A. A. WATSON, *The optimum distance at which to determine the size of a giant air shower*, *Astropart. Phys.*, 26 (2007), pp. 414–419.
- [61] M. ROTH, *Measurement of the UHECR energy spectrum using data from the Surface Detector of the Pierre Auger Observatory*, (2007).
- [62] C. BONIFAZI, A. LETESSIER-SELVON, AND E. SANTOS, *A model for the time uncertainty measurements in the Auger surface detector array*, *Astropart.Phys.*, 28 (2008), pp. 523–528.
- [63] J. HORANDEL *et al.*, *Test of the hadronic interaction model EPOS with KASCADE air shower data*, *Nucl.Phys.Proc.Suppl.*, 196 (2009), pp. 235–238.
- [64] M. LUDWIG AND T. HUEGE, *REAS3: Monte Carlo simulations of radio emission from cosmic ray air showers using an 'end-point' formalism*, *Astropart.Phys.*, 34 (2011), pp. 438–446.
- [65] J. A. ABRAHAM *et al.*, *Operations of and Future Plans for the Pierre Auger Observatory*, (2009).
- [66] J. C. DE NO, *Monitoring of camera positions for auger fluorescence telescopes*, master's thesis, RWTH Aachen University, 2008.
- [67] M. AVE *et al.*, *The 2pt+: an enhanced 2 point correlation function*, *JCAP*, 0907 (2009), p. 023.
- [68] R. BRUN AND F. RADEMAKERS, *ROOT: An object oriented data analysis framework*, *Nucl. Instrum. Meth.*, A389 (1997), pp. 81–86.
- [69] J. D. BARROW, S. P. BHAVSAR, AND D. H. SONODA, *Minimal spanning trees, filaments and galaxy clustering*, *Mon. Not. R. Astron. Soc.*, 216 (1985), pp. 17–35.
- [70] D. HARARI, S. MOLLERACH, AND E. ROULET, *Detecting filaments in the ultra-high energy cosmic ray distribution*, *Astropart. Phys.*, 25 (2006), pp. 412–418.
- [71] K. GORSKI *et al.*, *HEALPix - A Framework for high resolution discretization, and fast analysis of data distributed on the sphere*, *Astrophys.J.*, 622 (2005), pp. 759–771.
- [72] P. SOMMERS, *Cosmic Ray Anisotropy Analysis with a Full-Sky Observatory*, *Astropart. Phys.*, 14 (2001), pp. 271–286.
- [73] G. R. FARRAR, A. A. BERLIND, AND I. ZAW, *Correlations between Ultrahigh Energy Cosmic Rays and Infrared-Luminous Galaxies*, *Astrophys.J.*, 716 (2010), pp. 914–917.
- [74] G. NEUGEBAUER *et al.*, *The Infrared Astronomical Satellite (IRAS) mission*, *Astrophys.J.*, 278 (1984), pp. L1–L6.

- [75] W. SAUNDERS *et al.*, *The PSCz Catalogue*, Mon. Not. Roy. Astron. Soc., 317 (2000), p. 55.
- [76] R. D. BLANDFORD AND R. L. ZNAJEK, *Electromagnetic extractions of energy from Kerr black holes*, Mon. Not. Roy. Astron. Soc., 179 (1977), pp. 433–456.
- [77] H. FALCKE, *Black Hole, Jet, and Disk: The Universal Engine*, (1995).

Acknowledgements

Finally, it is time to recognise all those who helped and supported me over the last years. First of all I would like to thank Prof. Dr. Thomas Hebbeker for giving me the opportunity to participate in the Pierre Auger Observatory. He always showed me that there are still open questions to be answered and new things to learn. In addition, I would like to thank Prof. Dr. Martin Erdmann who agreed to be the co-corrector of this thesis.

Of course, I thank all former and current members of Auger group from Aachen for their steady support. Although some have a weird way of showing it. In alphabetical order my thanks goes to Hans Bretz, Joaquin Calvo de No, Dr. Hans Dembinski, Tim Enzweiler, Stefan Fliescher, Christoph Genreith, Marius Grigat, Markus Lauscher, Dr. Matthias Leuthold, Dr. Christine Meurer, Lukas Middendorf, Lars Mohrmann, Gero Müller, Anna Nelles, Tim Niggemann, Matthias Plum, Nils Scharf, Peter Schiffer, Sarah Schmetkamp, Johannes Schumacher, Oliver Seeger, Maurice Stephan, Marcel Straub, David Walz, Klaus Weidenhaupt and Tobias Winchen. Each one contributed in its own way to this thesis and made the time I spent working on it a real pleasure. I am very grateful for time I worked with all of them. In this context, I would also like to mention Jan Steggemann and Joschka Lingemann who spent nearly every lunch with us and enriched it with fruitful discussions on various topics.

For reading parts of this thesis, I give my special thanks to Maurice, Marius, Markus Merschmeyer, Nils and Martin Bissok. Even when they were pinched by time, they could find a spot to squeeze in a chapter of mine. Thanks a lot for your effort.

Furthermore, I need to thank the *Large Scale Anisotropy Task* and its task leaders Prof. Joao de Mello Neto and Dr. Olivier Deligny. I learned a lot during the sessions and corresponding discussions about anisotropies and how to find them. In this context, my special thanks goes to Haris Lyberis and Sarka Jiraskova, who were searching with me for the first point source of UHECRs.

I would like to thank Hans Klages, Steffen Müller and Nils for the help they gave me working on the TILT-monitoring system for HEAT. It would have taken much longer to install and to operate it alone.

A sweet thanks goes to Adriana del Piero, who supplied me every time with cookies when the need was greatest.

My deepest gratitude goes to Marius Grigat and Markus Merschmeyer. Both had always a sympathetic ear for my worries as well as problems and everything else that comes to mind. In this sense, the greater respect goes to Marius who suffered most sharing an office with me. Others would certainly have fled. I enjoyed it a lot to work, to discuss, to argue and to laugh with them. I will miss them, the Russian chocolate cheesecake and the sweet elderly couple.

Thanks goes to my family, who supported me during the years of my study without pushing me. They knew I needed my time.

Last but not least, I would like to add my best friends to this list who were always there for me, Martin Spinrath and Gunnar Bruns. With you guys I spent some of the best times in my life.

Lebenslauf

Persönliche Daten:

Nachname: Schulte
Vorname: Stephan
Geburtsdatum: 02.11.1981
Geburtsort: Arnsberg (Westfalen)
Staatsangehörigkeit: deutsch

Qualifikationen:

2001: Abitur
ab 2002 bis 2007: Studium der Physik, Abschluss Diplomphysiker
ab 2007: Beginn der Promotion